

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OW protein - protein search, using sw model

Run on: October 7, 2003, 07:04:18 : Search time 96 Seconds
(without alignments)
908.560 Million cell updates/sec

Title: US-09-856-679-2
Perfect score: 1759
Sequence: 1 MAALQEDGWTGQVLVKVN.....QLKVIDNQRELSRLELP 338

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

SPTREMBL_23: *
1: sp_archaea: *
2: sp_bacteria: *
3: sp_fungi: *
4: sp_human: *
5: sp_invertebrate: *
6: sp_mammal: *
7: sp_mhc: *
8: sp_organelle: *
9: sp_phage: *
10: sp_plant: *
11: sp_rodent: *
12: sp_virus: *
13: sp_vertebrate: *
14: sp_unclassified: *
15: sp_virus: *
16: sp_bacteriophage: *
17: sp_archaeophage: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Match Length	DB ID	Description
1	1759	100.0	881	4	Q8WVNO
2	1755	99.8	881	4	Q95398
3	1748	99.4	881	4	Q95634
4	1664	94.6	884	11	Q921C8
5	1655	94.1	461	11	Q8BZK9
6	1591	90.4	876	11	Q8VCC8
7	1432	80.8	287	11	Q8R1R1
8	863.5	49.1	1011	4	Q8WZAZ
9	863.5	49.1	1011	4	Q95636
10	857.5	48.7	1011	4	Q8TAA4
11	855.5	48.6	696	11	Q8VIRP
12	855.5	48.6	699	11	Q9CWS2
13	855.5	48.6	867	4	Q81XK6
14	855.5	48.6	993	11	Q921P0
15	853.5	48.6	1011	11	Q9E0Z6
16	832.5	47.3	436	11	Q921C7

17	696	39.6	861	5	Q9V9A3	Q9V9A3 drosophila
18	655	37.2	580	4	Q92565	Q92565 homo sapien
19	644.5	36.6	612	11	Q8B1J9	Q8B1J9 mus musculu
20	644.5	36.6	814	11	Q8C0Q9	Q8C0Q9 mus musculu
21	637.5	36.2	456	4	Q90HV5	Q90HV5 homo sapien
22	637.5	36.2	814	11	Q8C0R5	Q8C0R5 mus musculu
23	495.5	28.2	444	4	Q81XU5	Q81XU5 homo sapien
24	459	26.1	1311	5	Q95NL8	Q95NL8 caenorhabd
25	459	26.1	1347	5	Q95WR8	Q95WR8 caenorhabd
26	459	26.1	1470	5	Q212I8	Q212I8 caenorhabd
27	409	23.3	1499	4	Q914G8	Q914G8 homo sapien
28	406	23.1	1138	11	Q8CHG7	Q8CHG7 mus musculu
29	404.5	23.0	1573	5	Q95V18	Q95V18 drosophila
30	402.5	22.9	1573	5	Q9VWF3	Q9VWF3 drosophila
31	390.5	22.2	1204	4	Q9UHV4	Q9UHV4 homo sapien
32	390.5	22.2	1391	4	Q8TEU6	Q8TEU6 homo sapien
33	390.5	22.2	1601	4	Q8TEU7	Q8TEU7 homo sapien
34	390.5	22.2	1601	4	Q8N121	Q8N121 homo sapien
35	387.5	22.0	834	11	Q8R3E5	Q8R3E5 mus musculu
36	386.5	22.0	1509	4	Q96PC1	Q96PC1 homo sapien
37	380.5	21.6	1113	4	Q8REA3	Q8REA3 homo sapien
38	306.5	17.4	929	5	Q8SSQ3	Q8SSQ3 dicyostell
39	304.5	17.3	1721	5	Q8SSQ0	Q8SSQ0 dicyostell
40	301.5	17.1	824	5	Q8IS15	Q8IS15 dicyostell
41	301	17.1	812	5	Q8IS14	Q8IS14 dicyostell
42	283.5	16.1	642	11	Q8C5V7	Q8C5V7 mus musculu
43	283.5	16.1	1086	11	Q91Z22	Q91Z22 mus musculu
44	283	16.1	1095	4	Q8IV73	Q8IV73 homo sapien
45	279	15.9	1013	11	Q9QYV3	Q9QYV3 rattus norv

ALIGNMENTS

RESULT 1
Q8WVNO PRELIMINARY; PRT; 881 AA.
ID Q8WVNO
AC Q8WVNO;
DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 21, Last annotation update)
DE 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Rapi guanine-nucleotide-exchange factor directly activated by
DE CAMP.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN (1)
RP SEQUENCE FROM N.A.
RA TISSUE=Kidney;
RA Strausberg R.;
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
EMBL; BC017728; AAH17728.1; -
DR InterPro: IPR002373; CAMP_kin.
DR InterPro: IPR000595; CNMP_binding.
DR InterPro: IPR000591; DEP.
DR InterPro: IPR000651; RASGEF.
DR InterPro: IPR001895; RASGEF_CDC25.
DR Pfam: PF00027; CNMP_binding; 1.
DR Pfam: PF00610; DEP; 1.
DR Pfam: PF00617; RASGEF; 1.
DR Pfam: PF00618; RASGEF; 1.
DR PRINTS: PR00103; CAMPKINASE.
DR SMART: SM00100; CNMP; 1.
DR SMART: SM00449; DEP; 1.
DR SMART: SM00147; RASGEF; 1.
DR SMART: SM00229; RASGEF; 1.
DR PROSITE: PS00042; CNMP_BINDING_3; 1.
DR PROSITE: PS00186; DEP; 1.
SQ SEQUENCE 881 AA; 99352 MW; D1045DZADPFB29A CRC64;

Query Match 100.0%; Score 1759; DB 4; Length 881;
Best Local Similarity 100.0%; Pred. No. 5e-152;

```
Matches 338; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MAALAQEDGWTGGVLYKVNNSAGDAIGLQDPARGVATSLGNERLFVYNPOEVHLLIHP 60
DB 544 MAALAQEDGWTGGVLYKVNNSAGDAIGLQDPARGVATSLGNERLFVYNPOEVHLLIHP 603
QY 61 DQGPVTGASAGLDLVSAKDLAAGLTBDHMSLFNSIHQVELIHVVLGPQHLRDVTYANLE 120
DB 604 DQGPVTGASAGLDLVSAKDLAAGLTBDHMSLFNSIHQVELIHVVLGPQHLRDVTYANLE 663
QY 121 RFMRFRNELQYVWATELCLCPVPGPRAQLRKFTKLAHLKEOKNLSFFAVMGLSNSA 180
DB 664 RFMRFRNELQYVWATELCLCPVPGPRAQLRKFTKLAHLKEOKNLSFFAVMGLSNSA 723
QY 181 ISRLAHTWERLPHVVRKLYSALERLDPSPMNHRYRLALAKLSPVPIPFMPLLLKDMTFI 240
DB 724 ISRLAHTWERLPHVVRKLYSALERLDPSPMNHRYRLALAKLSPVPIPFMPLLLKDMTFI 783
QY 241 HEGNHTLVENLINFEKRMMAARAAMLHCHRSNHPVPLSPRSVSHLHEDSOVARISTC 300
DB 784 HEGNHTLVENLINFEKRMMAARAAMLHCHRSNHPVPLSPRSVSHLHEDSOVARISTC 843
QY 301 SEQSLSTRSPASTWAVYQOLKVIDNORELSRLSRELEP 338
DB 844 SEQSLSTRSPASTWAVYQOLKVIDNORELSRLSRELEP 881

RESULT 2
095398 PRELIMINARY; PRT; 881 AA.
AC 095398;
DT 01-MAY-1999 (TREMBLrel. 10, Created)
DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)
DE 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE Rap1 guanine-nucleotide exchange factor.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Muscle;
RX MEDLINE=99066645; PubMed=9853756;
RA de Rooij J., Zwartkruis F.J., Verheijen M.H., Cool R.H., Nijman S.M.,
RA Wittinghofer A., Bos J.L.;
RT "Epac is a Rap1 guanine-nucleotide-exchange factor directly activated
RT by cyclic AMP.";
RL Nature 396:474-477(1998).
DR EMBL; AF103905; AAC83381.1; -
DR InterPro; IPR002373; CAMP_kin.
DR InterPro; IPR000595; CNMP_binding.
DR InterPro; IPR000651; DEP.
DR InterPro; IPR001895; RasGEF.
DR pfam; PF00027; CNMP_binding; 1.
DR pfam; PF00610; DEP; 1.
DR pfam; PF00617; RasGEF; 1.
DR PRINTS; PRO0103; CAMPKINASE.
DR SMART; SM00100; CNMP; 1.
DR SMART; SM00049; DEP; 1.
DR SMART; SM00147; RasGEF; 1.
DR SMART; SM00229; RasGEFN; 1.
DR PROSITE; PS50042; CNMP_BINDING_3; 1.
DR PROSITE; PS50186; DEP; 1.
SQ SEQUENCE 881 AA; 99312 MW; 76B82C54E0316D45 CRC64;

Query Match 99.8%; Score 1755; DB 4; Length 881;
Best Local Similarity 99.7%; Pred. No. 1.2e-151;
Matches 337; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 MAALAQEDGWTGGVLYKVNNSAGDAIGLQDPARGVATSLGNERLFVYNPOEVHLLIHP 60
DB 544 MAALAQEDGWTGGVLYKVNNSAGDAIGLQDPARGVATSLGNERLFVYNPOEVHLLIHP 603
```

```
DB 544 MAALAQEDGWTGGVLYKVNNSAGDAIGLQDPARGVATSLGNERLFVYNPOEVHLLIHP 603
QY 61 DQGPVTGASAGLDLVSAKDLAAGLTBDHMSLFNSIHQVELIHVVLGPQHLRDVTYANLE 120
DB 604 DQGPVTGASAGLDLVSAKDLAAGLTBDHMSLFNSIHQVELIHVVLGPQHLRDVTYANLE 663
QY 121 RFMRFRNELQYVWATELCLCPVPGPRAQLRKFTKLAHLKEOKNLSFFAVMGLSNSA 180
DB 664 RFMRFRNELQYVWATELCLCPVPGPRAQLRKFTKLAHLKEOKNLSFFAVMGLSNSA 723
QY 181 ISRLAHTWERLPHVVRKLYSALERLDPSPMNHRYRLALAKLSPVPIPFMPLLLKDMTFI 240
DB 724 ISRLAHTWERLPHVVRKLYSALERLDPSPMNHRYRLALAKLSPVPIPFMPLLLKDMTFI 783
QY 241 HEGNHTLVENLINFEKRMMAARAAMLHCHRSNHPVPLSPRSVSHLHEDSOVARISTC 300
DB 784 HEGNHTLVENLINFEKRMMAARAAMLHCHRSNHPVPLSPRSVSHLHEDSOVARISTC 843
QY 301 SEQSLSTRSPASTWAVYQOLKVIDNORELSRLSRELEP 338
DB 844 SEQSLSTRSPASTWAVYQOLKVIDNORELSRLSRELEP 881

RESULT 3
095634 PRELIMINARY; PRT; 881 AA.
AC 095634;
DT 01-MAY-1999 (TREMBLrel. 10, Created)
DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)
DE 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE CAMP-regulated guanine nucleotide exchange factor I.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99074384; PubMed=9856955;
RA Kawasaki H., Springett G.M., Mochizuki N., Toki S., Nakaya M.,
RA Matsuda M., Housman D.E., Graybiel A.M.;
RT "A family of CAMP-binding proteins that directly activate rap1.";
RL Science 282:2275-2279(1998).
DR EMBL; U78168; AAD12740.1; -
DR InterPro; IPR000595; CNMP_binding.
DR InterPro; IPR000651; DEP.
DR InterPro; IPR001895; RasGEF.
DR pfam; PF00027; CNMP_binding; 1.
DR pfam; PF00610; DEP; 1.
DR pfam; PF00617; RasGEF; 1.
DR PRINTS; PRO0103; CAMPKINASE.
DR SMART; SM00100; CNMP; 1.
DR SMART; SM00049; DEP; 1.
DR SMART; SM00147; RasGEF; 1.
DR SMART; SM00229; RasGEFN; 1.
DR PROSITE; PS50042; CNMP_BINDING_3; 1.
DR PROSITE; PS50186; DEP; 1.
SQ SEQUENCE 881 AA; 99376 MW; 971CA7A6EEF7BF1E CRC64;

Query Match 99.4%; Score 1748; DB 4; Length 881;
Best Local Similarity 99.1%; Pred. No. 5.1e-151;
Matches 335; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
QY 1 MAALAQEDGWTGGVLYKVNNSAGDAIGLQDPARGVATSLGNERLFVYNPOEVHLLIHP 60
DB 544 MAALAQEDGWTGGVLYKVNNSAGDAIGLQDPARGVATSLGNERLFVYNPOEVHLLIHP 603
QY 61 DQGPVTGASAGLDLVSAKDLAAGLTBDHMSLFNSIHQVELIHVVLGPQHLRDVTYANLE 120
DB 604 DQGPVTGASAGLDLVSAKDLAAGLTBDHMSLFNSIHQVELIHVVLGPQHLRDVTYANLE 663
QY 121 RFMRFRNELQYVWATELCLCPVPGPRAQLRKFTKLAHLKEOKNLSFFAVMGLSNSA 180
```

```
Db 664 RFRNREFNELOQWVATLCLCPVPGPRAQLKKFKIKLAHLKEQKNVSPFAVMGSLNSP 723
Qy 181 ISRLAHTWBERLPHKVRKLYSALERLDDPSNNHRYRLATKLSPVPIPFMPLLLKDMFTI 240
Db 724 ISRLAHTWBERLPHKVRKLYSALERLDDPSNNHRYRLATKLSPVPIPFMPLLLKDMFTI 763
Qy 241 HEGNHTLVENLINFCKRMMAARAARMLHCRSHNPVPLSPRSVSHLHDSQVARIStC 300
Db 784 HEGNHTLVENLINFCKRMMAARAARMLHCRSHNPVPLSPRSVSHLHDSQVARIStC 843
Qy 301 SEQSLSTRSPASTWAVYQOLKVIDNORELSRLSRELEP 338
Db 844 SEQSLSTRSPASTWAVYQOLKVIDNORELSRLSRELEP 881

RESULT 4
ID 0921C8 PRELIMINARY; PRT: 884 AA.
AC 0921C8;
DT 01-MAY-1999 (TReMBLrel. 10, Created)
DT 01-MAY-1999 (TReMBLrel. 10, Last sequence update)
DE 01-MAR-2003 (TReMBLrel. 23, Last annotation update)
DE CAMP-regulated guanine nucleotide exchange factor 1.
GN CAMP-GEFI.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99074384; Pubmed=9856955;
RA Kawasaki H., Sprioggett G.M., Mochizuki N., Toki S., Nakaya M.,
RA Metada M., Housman D.E., Graybiel A.M.,
RT "A family of CAMP-binding proteins that directly activate rap1."
RL Science 282:2275-2279(1998).
EMBL: U78167; AAD12739.1; -.
DR InterPro: IPR002373; CAMP_kin.
DR InterPro: IPR000595; CAMP_binding.
DR InterPro: IPR000591; DEP.
DR InterPro: IPR000651; RAGEFN.
DR InterPro: IPR001895; RAGEFN_CDC25.
DR Pfam: PF00027; CNMP_binding.1.
DR Pfam: PF00610; DEP.1.
DR Pfam: PF00617; RAGEFN.1.
DR Pfam: PF00618; RAGEFN.1.
DR PRINTS; PRO0103; CAMPKINASE.
DR SMART; SM00100; CNMP.1.
DR SMART; SM00049; DEP.1.
DR SMART; SM00147; RAGEFN.1.
DR SMART; SM00229; RAGEFN.1.
DR PROSITE; PSS0042; CNMP_BINDING_3.1.
DR PROSITE; PSS0186; DEP.1.
SQ SEQUENCE 884 AA; 100256 MW; B082AAE3155AC933 CRC64;

Query Match 94.6%; Score 1664; DB 11; Length 884;
Best Local Similarity 93.5%; Pred. No. 2,5e-143;
Matches 316; Conservative 11; Mismatches 11; Indels 0; Gaps 0;

Qy 1 MAALAEEDGWTGKGOVLKVNVSAGDAIGLPDARGVATSLGILNRLFYVNPQEVHELIRHP 60
Db 547 MAALAEEDGWTGKGOVLKVNVSAGDAIGLPDARGVATSLGILNRLFYVNPQEVHELIRHP 606
Qy 61 DQGLPTVGSABGLDVSARKLAQGLTDHMSLPSNSTHIOVELINHYVLGPOHLRDVTTANLE 120
Db 607 DQGLPTVGSABGLDVSARKLAQGLTDHMSLPSNSTHIOVELINHYVLGPOHLRDVTTANLE 666
Qy 121 RFRNREFNELOQWVATLCLCPVPGPRAQLKKFKIKLAHLKEQKNLSFPAVMGSLNSA 180
Db 667 RFRNREFNELOQWVATLCLCPVPGPRAQLKKFKIKLAHLKEQKNLSFPAVMGSLNSA 726
Qy 181 ISRLAHTWBERLPHKVRKLYSALERLDDPSNNHRYRLATKLSPVPIPFMPLLLKDMFTI 240
```

```
Db 727 ISRLAHTWBERLPHKVRKLYSALERLDDPSNNHRYRLATKLSPVPIPFMPLLLKDMFTI 786
Qy 241 HEGNHTLVENLINFCKRMMAARAARMLHCRSHNPVPLSPRSVSHLHDSQVARIStC 300
Db 787 HEGNHTLVENLINFCKRMMAARAARMLHCRSHNPVPLSPRSVSHLHDSQVARIStC 846
Qy 301 SEQSLSTRSPASTWAVYQOLKVIDNORELSRLSRELEP 338
Db 847 SEQSLSTRSPASTWAVYQOLKVIDNORELSRLSRELEP 884

RESULT 5
ID 08B2K9 PRELIMINARY; PRT: 461 AA.
AC 08B2K9;
DT 01-MAR-2003 (TReMBLrel. 23, Created)
DT 01-MAR-2003 (TReMBLrel. 23, Last sequence update)
DE 01-MAR-2003 (TReMBLrel. 23, Last annotation update)
DE CAMP-regulated guanine nucleotide exchange factor I homolog
DE (Fragment).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=C57BL/6J; TISSUE=Dienecephalon;
RX MEDLINE=22354683; Pubmed=12466851;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs."
RL Nature 420:563-573(2002).
EMBL: AK034265; BAC28653.1; -.
FT NON_CODING
SQ SEQUENCE 461 AA; 52434 MW; 638101E1D2E4249 CRC64;

Query Match 94.1%; Score 1655; DB 11; Length 461;
Best Local Similarity 93.5%; Pred. No. 6.8e-143;
Matches 316; Conservative 10; Mismatches 12; Indels 0; Gaps 0;

Qy 1 MAALAEEDGWTGKGOVLKVNVSAGDAIGLPDARGVATSLGILNRLFYVNPQEVHELIRHP 60
Db 124 MAALAEEDGWTGKGOVLKVNVSAGDAIGLPDARGVATSLGILNRLFYVNPQEVHELIRHP 183
Qy 61 DQGLPTVGSABGLDVSARKLAQGLTDHMSLPSNSTHIOVELINHYVLGPOHLRDVTTANLE 120
Db 184 DQGLPTVGSABGLDVSARKLAQGLTDHMSLPSNSTHIOVELINHYVLGPOHLRDVTTANLE 243
Qy 121 RFRNREFNELOQWVATLCLCPVPGPRAQLKKFKIKLAHLKEQKNLSFPAVMGSLNSA 180
Db 244 RFRNREFNELOQWVATLCLCPVPGPRAQLKKFKIKLAHLKEQKNLSFPAVMGSLNSA 303
Qy 181 ISRLAHTWBERLPHKVRKLYSALERLDDPSNNHRYRLATKLSPVPIPFMPLLLKDMFTI 240
Db 304 ISRLAHTWBERLPHKVRKLYSALERLDDPSNNHRYRLATKLSPVPIPFMPLLLKDMFTI 363
Qy 241 HEGNHTLVENLINFCKRMMAARAARMLHCRSHNPVPLSPRSVSHLHDSQVARIStC 300
Db 364 HEGNHTLVENLINFCKRMMAARAARMLHCRSHNPVPLSPRSVSHLHDSQVARIStC 423
Qy 301 SEQSLSTRSPASTWAVYQOLKVIDNORELSRLSRELEP 338
Db 424 SEQSLSTRSPASTWAVYQOLKVIDNORELSRLSRELEP 461

RESULT 6
ID 08VCC8 PRELIMINARY; PRT: 876 AA.
AC 08VCC8;
DT 01-MAR-2002 (TReMBLrel. 20, Created)
DT 01-MAR-2002 (TReMBLrel. 20, Last sequence update)
DT 01-OCT-2002 (TReMBLrel. 22, Last annotation update)
```

DE Similar to cAMP-regulated guanine nucleotide exchange factor I (cAMP-
DE GEF1).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RA Strausberg R.;
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: BC020533; AAH20533.1; -
DR InterPro: IPR002373; CAMP_kin
DR InterPro: IPR000595; CAMP_binding.
DR InterPro: IPR000591; DEP.
DR InterPro: IPR000651; RASGEFN.
DR InterPro: IPR001895; RASGRF_CCC25.
DR Pfam: PF00027; CNMP_binding; 1.
DR Pfam: PF00610; DEP; 1.
DR Pfam: PF00617; RASGEF; 1.
DR Pfam: PF00618; RASGEFN; 1.
DR PRINTS: PR00103; CAMPKINASE.
DR SMART: SM00100; CNMP; 1.
DR SMART: SM00049; DEP; 1.
DR SMART: SM00147; RASGEF; 1.
DR SMART: SM00229; RASGEFN; 1.
DR PROSITE: PS50042; CNMP_BINDING_3; 1.
DR PROSITE: PS50186; DEP; 1.
SQ SEQUENCE: 876 AA; 99230 MW; 7DC12B83F18F9AF8 CRC64;

Query Match 90.4%; Score 1591; DB 11; Length 876;
Best Local Similarity 90.5%; Pred. No. 1.2e-136;
Matches 306; Conservative 12; Mismatches 12; Indels 8; Gaps 1;

QY 1 MAALAEQGWTKGQVLYVNSAGDAIGLPDARGVATSLGNERLFFVNPQEVHELIPHP 60
DB 547 MAALAEHDMTKGQVLYVNSAGDVYGLQPDARGVATSLGNERLFFVNPQEVHELIPHP 606
QY 61 DQLGPTVGSAGDLVSAKDLAAGLTLDHWSLFNSIHQVELIHVYLGQHLDTVTANLE 120
DB 607 EQLGPTLSSEMLDLVSAKDLAAGLTLDHDMNLFNRHQVQ-----EHLADVTYANLE 658
QY 121 RMRKRFNELQVVAATELCPCVPGRRAQLLRKFTLAHLKEQKLNSEFFAVMFLSNSA 180
DB 659 RMRKRFNELQVVAATELCPCVPGRRAQLLRKFTLAHLKEQKLNSEFFAVMFLSNSA 718
QY 181 ISRLAHTWERLPHKRYKLYSALERLLDPSMNRVRYRLALAKLSPVIFPMPLLDKMTFI 240
DB 719 ISRLAHTWERLPHKRYKLYSALERLLDPSMNRVRYRLALAKLSPVIFPMPLLDKMTFI 778
QY 241 HEGNHTLVENLINFKRMMAARAAMLHCRSHNFPVPLSRVSHLHEDSQVARISTC 300
DB 779 HEGNHTLVENLINFKRMMAARAAMLHCRSHNFPVPLSRVSHLHEDSQVARISTC 838
QY 301 SEQSISTSPASTAWYVOQLKYIDNORELSRLSRELEP 338
DB 839 SEQSISTSPASTAWYVOQLKYIDNORELSRLSRELEP 876

RESULT 7

Q8R1R1 PRELIMINARY; PRT; 287 AA.
AC Q8R1R1;
DT 01-JUN-2002 (TREMBLrel. 21, Created)
DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)
DT 01-OCT-2002 (TREMBLrel. 22, Last annotation update)
DE Hypothetical 33.2 kDa protein (Fragment).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.

RC TISSUE=Kidney;
RA Strausberg R.;
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: BC020311; AAH20311.1; -
DR InterPro: IPR001895; RASGRF_CDC25.
DR Pfam: PF00617; RASGEF; 1.
DR SMART: SM00147; RASGEF; 1.
KM Hypothetical protein.
FT NON_TER
SQ SEQUENCE 287 AA; 33165 MW; 87D9AF9505928C4B CRC64;

Query Match 80.8%; Score 1422; DB 11; Length 287;
Best Local Similarity 94.1%; Pred. No. 7.6e-122;
Matches 270; Conservative 8; Mismatches 9; Indels 0; Gaps 0;

QY 52 EVHELIPHPDQLGPTVGSAGDLVSAKDLAAGLTLDHWSLFNSIHQVELIHVYLGQHL 111
DB 1 EVHELIPHPDQLGPTVGSAGDLVSAKDLAAGLTLDHDMNLFNRHQVQHL 60
QY 112 RDVTANLERMRKRFNELQVVAATELCPCVPGRRAQLLRKFTLAHLKEQKLNSEFFA 171
DB 61 RDVTANLERMRKRFNELQVVAATELCPCVPGRRAQLLRKFTLAHLKEQKLNSEFFA 120
QY 172 VMFGLNSAISRLAHTWERLPHKRYKLYSALERLLDPSMNRVRYRLALAKLSPVIFPM 231
DB 121 VMFGLNSAISRLAHTWERLPHKRYKLYSALERLLDPSMNRVRYRLALAKLSPVIFPM 180
QY 232 LLLKDMFTIHEGNHTLVENLINFKRMMAARAAMLHCRSHNFPVPLSRVSHLHED 291
DB 181 LLLKDMFTIHEGNHTLVENLINFKRMMAARAAMLHCRSHNFPVPLSRVSHLHED 240
QY 292 SQVARISTCSQSISTSPASTAWYVOQLKYIDNORELSRLSRELEP 338
DB 241 SQGSRISTCSQSISTSPASTAWYVOQLKYIDNORELSRLSRELEP 287

RESULT 8

Q8WZ22 PRELIMINARY; PRT; 1011 AA.
AC Q8WZ22;
DT 01-MAR-2002 (TREMBLrel. 20, Created)
DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE CAMP-GEFI.
GN Homo sapiens (Human).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21564207; PubMed=11707077;
RA Ueno H., Shibasaki T., Iwanaga T., Takahashi K., Yokoyama Y.,
RA Liu L.N., Yokoi N., Ozaki N., Matsukura S., Yano H., Saito S.;
RT "Characterization of the gene EPAC2: structure, chromosomal
RT localization, tissue expression, and identification of the liver-
RT specific isoform";
RL Genomics 78:91-98 (2001).
DR EMBL: AB027471; BAB72179.1; -
DR InterPro: IPR000595; CNMP_binding.
DR InterPro: IPR000591; DEP.
DR InterPro: IPR000651; RASGEFN.
DR InterPro: IPR001895; RASGRF_CCC25.
DR Pfam: PF00027; CNMP_binding; 2.
DR Pfam: PF00610; DEP; 1.
DR Pfam: PF00617; RASGEF; 1.
DR Pfam: PF00618; RASGEFN; 1.
DR SMART: SM00100; CNMP; 2.
DR SMART: SM00049; DEP; 1.
DR SMART: SM00147; RASGEF; 1.
DR SMART: SM00229; RASGEFN; 1.
DR PROSITE: PS50042; CNMP_BINDING_3; 2.
DR PROSITE: PS50186; DEP; 1.

Seq	Sequence	1011 AA	115521 MW	9427C5P92F2FFAD1 CRC64
QY	Query Match	49.1%	Score 863.5	DB 4
Db	Best Local Similarity	50.0%	Pred. No. 5.8e-70	
Matches	169	Conservative	63	Mismatches 81; Indels 25; Gaps 3
QY	1	MAALAEQGMKRGQVLVKNVNSAGDILGQDPARGVATSLGILNERLFVYVNPQGVHLLIRP	60	
Db	696	ISAAVADKSGEGGLIVMSSGGEVYVHLKPNDSVFTLLTNGRLFLACPRREGDSLTPLP	755	
QY	61	DOLGPVTSAGGLDLSAKDLAGOLQTDIDMSLFSNIHOVELHYVLCPOHLDVYTTANLE	120	
Db	756	EQEGPTVGTCTFELMSSKDLAQYQTTITDWEFLNFCVHELELITHTFG-RHNKKTNTAND	814	
QY	121	RPMRRFNLQVWATVETLCLCPVPGPRAOLLKRFIKLAHLKEQKNLSFFAVMGLSNSA	180	
Db	815	LFLRFNFIEIOFWVTEICLSQLSRVOLLKKFKIFIAHCKEYKNLNSFFAIVGLSNVA	874	
QY	181	ISRLAHMTERLPHKVRKILYSALERLDDSDMNRVRYRLAKSPVPIPPMLLDMMFTI	240	
Db	875	VSRLALTEWELPSKRRKRYAEESLMDPSRNRRAVRLVYAKLEPILIPMPLIIDMMFT	934	
QY	241	HEGNTLVEITLNFEEKMMARAAMLHHCRSHNPVPSLRVSHLHEDSQVARISFC	300	
Db	935	HEGKNTFDNLNVEPKMKRMINTATATVYYSQ---PKNPDAQAANKKHQDV-----	984	
QY	301	SEQSILSTRSPASTVAVYVOOLVVIDNQELSRLSRELP	338	
Db	985	-----SYVRLNVIDNQRLSQSHRLEP	1008	
RESULT 9				
QY	095636	PRELIMINARY	PRT	1011 AA
ID	095636			
AC	095636			
DT	01-MAY-1999	(TREMBLrel, 10, Created)		
DT	01-MAY-1999	(TREMBLrel, 10, Last sequence update)		
DE	01-MAR-2003	(TREMBLrel, 23, Last annotation update)		
DE	CAMP-regulated	guanine nucleotide exchange factor II.		
GN	CAMP-GEFII.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
CC	Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.			
OX	NCBI_Taxid=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=96074384; PubMed=9856955;			
RA	Kawasaki H., Springett G.M., Mochizuki N., Toki S., Nakaya M.,			
RA	Matsuda M., Housman D.E., Graybiel A.M.;			
RT	"A family of CAMP-binding proteins that directly activate rap1,"			
RL	Science 282:2275-2279(1998).			
DR	EMBL; U78516; AAD03422.1;			
DR	InterPro: IPR000595; CNMP_binding.			
DR	InterPro: IPR000591; DEP.			
DR	InterPro: IPR000651; RasGEFN.			
DR	InterPro: IPR001893; RasGRF_CDC25.			
DR	Pfam: PF00610; DEP; 1.			
DR	Pfam: PF00617; RasGEF; 1.			
DR	Pfam: PF00618; RasGEFN; 1.			
DR	SMART; SM00100; CNMP; 2.			
DR	SMART; SM0049; DEP; 1.			
DR	SMART; SM00147; RasGEF; 1.			
DR	SMART; SM00229; RasGEFN; 1.			
DR	PROSITE: PS50042; CNMP_BINDING_3; 2.			
DR	PROSITE: PS50186; DEP; 1.			
SO	SEQUENCE	1011 AA; 115535 MW; 824B1B36E78CFAD1 CRC64;		
QY	Query Match	49.1%	Score 863.5	DB 4
Db	Best Local Similarity	50.0%	Pred. No. 5.8e-70	
Matches	169	Conservative	63	Mismatches 81; Indels 25; Gaps 3
QY	1	MAALAEQGMKRGQVLVKNVNSAGDILGQDPARGVATSLGILNERLFVYVNPQGVHLLIRP	60	

DB	Query Match	Similarity	Score	DB 4	Length	1011
Db	636	ISAVADKLGSGEGLIIVKMSSGGEVYVLLKPNDSVFTTLITIGRLFPACPREDFDSLTPLP	48.7%	49.7%	Pred No. 2.1e-69;	755
Qy	61	DOLGPTVGSAGLGLDVSARKDLAAGULTDHDMSLFNSIHQVELIHYVGLGPHLNDVTTANLE				120
Db	756	EOEGTGTGTGVEFELMSSKDLAYQNTIYDWELEFNCVHELELLYHFFG-RHNKKTGTAND				814
Qy	121	RPMRRENLQYVATVELCLCPVPGRFQAOLLRKFIFYLAHLKEOKMLNSFFAVMFGLSNSA				180
Db	815	LFLRRFENIQYVWVTEICLSOLSKRYVOLLKKFKIFYIAHCKEYKNLNSFFAIVMGLSNIA				874
Qy	181	ISRLAHTMERLPKRYKRLYSALERLDDPSMNRVYRLALAKSPVIEPMPILLDMFPI				240
Db	875	VSRLATLWEKLSKPKKTYAEBSLMDPSRRNRAVRLVYAKLEPPLIPMPILLDMFTT				934
Qy	241	HEGNNTFTLVNLFNEFKMKMINTAFTATYRYSQ---PFPDPAQAQANKNHQDVR-----				300
Db	935	HEGNNTFTLVNLFNEFKMKMINTAFTATYRYSQ---PFPDPAQAQANKNHQDVR-----				984
Qy	301	SEQSILSTRSPASTWAVYQOLKVIDNORELSRLSLEP				338
Db	985	-----SYVQLNVINDQRTLSQMSHRLP				1008
RESULT 10						
ID	Q8TAA4	PRELIMINARY;	PRT;	1011	AA.	
AC	Q8TAA4;					
DT	01-JUN-2002 (TREMBlrel. 21, Created)					
DT	01-JUN-2002 (TREMBlrel. 21, Last sequence update)					
DT	01-MAR-2003 (TREMBlrel. 23, Last annotation update)					
DE	CAMP-regulated guanine nucleotide exchange factor II.					
OS	Homo sapiens (Human).					
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;					
OC	Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.					
OX	NCBI_TaxID=9606;					
RN	[1]					
RP	SEQUENCE FROM N.A.					
RC	TISSUE=Testis;					
RA	Strausberg R.;					
RL	Submitted (FEB-2002) to the EMBL/Genbank/DBJ databases.					
DR	EMBL: BC024004; AAH24004.1; -					
DR	InterPro: IPR000595; CNMP_binding.					
DR	InterPro: IPR000591; DEP.					
DR	InterPro: IPR001895; RasGEFN.					
DR	InterPro: IPR00027; CNMP_binding; 2.					
DR	Pfam: PF00610; DEP; 1.					
DR	Pfam: PF00617; RasGEF; 1.					
DR	Pfam: PF00618; RasGEFN; 1.					
DR	SMART: SM00100; CNMP; 2.					
DR	SMART: SM00049; DEP; 1.					
DR	SMART: SM00147; RasGEF; 1.					
DR	SMART: SM00229; RasGEFN; 1.					
DR	PROSITE: PS50042; CNMP_BINDING_3; 2.					
DR	PROSITE: PS50186; DEP; 1.					
SO	SEQUENCE	1011	AA;	115511	MM;	910EA3651CC9B34A CRC64;
Query Match						
Db	696	ISAVADKLGSGEGLIIVKMSSGGEVYVLLKPNDSVFTTLITIGRLFPACPREDFDSLTPLP				755
Qy	1	MAALAQEDGWTGQVLYLVNSAGDAIGLQDPARGVATSIGLERLFVNVNPOEHLIPHP				60
Db	696	ISAVADKLGSGEGLIIVKMSSGGEVYVLLKPNDSVFTTLITIGRLFPACPREDFDSLTPLP				755
Qy	61	DOLGPTVGSAGLGLDVSARKDLAAGULTDHDMSLFNSIHQVELIHYVGLGPHLNDVTTANLE				120
Db	756	EOEGTGTGTGVEFELMSSKDLAYQNTIYDWELEFNCVHELELLYHFFG-RHNKKTGTAND				814
Qy	121	RPMRRENLQYVATVELCLCPVPGRFQAOLLRKFIFYLAHLKEOKMLNSFFAVMFGLSNSA				180
Db	815	LFLRRFENIQYVWVTEICLSOLSKRYVOLLKKFKIFYIAHCKEYKNLNSFFAIVMGLSNIA				874
Qy	181	ISRLAHTMERLPKRYKRLYSALERLDDPSMNRVYRLALAKSPVIEPMPILLDMFPI				240
Db	875	VSRLATLWEKLSKPKKTYAEBSLMDPSRRNRAVRLVYAKLEPPLIPMPILLDMFTT				934
Qy	241	HEGNNTFTLVNLFNEFKMKMINTAFTATYRYSQ---PFPDPAQAQANKNHQDVR-----				300
Db	935	HEGNNTFTLVNLFNEFKMKMINTAFTATYRYSQ---PFPDPAQAQANKNHQDVR-----				984
Qy	301	SEQSILSTRSPASTWAVYQOLKVIDNORELSRLSLEP				338
Db	985	-----SYVQLNVINDQRTLSQMSHRLP				1008

Qy	181	ISRLAHMERLPHKYRKLXLSALERLDDSMNHRVRLAKLSPIVFPMLLKDMTFI	240
Db	875	VSRLALTEKRLPSFKFKYAEFESLMDSSRNHRATRLVAKLEPLIFPMLLIDMTFT	934
Qy	241	HEGNHTVENLINFEMKRMARAAEMLHNCRSHNPPLSPKRSVSHLEDSQVARI	300
Db	935	HEGNKTFIDNLVNEFKMKMIANTARTVRYRSQ---PPNPDAQANKNHQDVR-----	984
Qy	301	SEQSISTSPASTAAVYQOLKAVINDQRLSLRSLRLEP	338
Db	985	-----SYROLANVIDNQRLSLQMSHRLP	1008
RESULT 11			
Q8V1P9	PRELIMINARY:	PRT:	696 AA.
AC	Q8V1P9:		
DT	01-MAR-2002 (TREMBLrel. 20, Created)		
DT	01-MAR-2002 (TREMBLrel. 20, Last sequence update)		
DT	01-OCT-2002 (TREMBLrel. 22, Last annotation update)		
DE	CAMP-GEFII (Fragment).		
GN	CGER2 OR EPAC2.		
OS	Mus musculus (Mouse).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
CC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
OX	NCBI_TaxId=10090;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	TISSUE=Liver;		
RX	MEDLINE=21564207; PubMed=11707077;		
RA	Deno H., Shibasaki T., Iwanaga T., Takahashi K., Yokoyama Y.,		
RA	Liu L.M., Yokoi N., Ozaki N., Matsukura S., Yano H., Saito S.;		
RT	*Characterization of the gene EPAC2: structure, chromosomal		
RT	localization, tissue expression, and identification of the liver-		
RT	specific isoform.;		
RL	Genomics 78:91-98(2001).		
DR	EMBL; AB037668; BAB72180.1; -.		
DR	MGI; MGI:1917723; Cgef2.		
DR	InterPro: IPR002373; CAMP_kin.		
DR	InterPro: IPR000595; CAMP_binding.		
DR	InterPro: IPR000651; RASGEFN.		
DR	InterPro: IPR01895; RASGRF_CDC25.		
DR	Pfam: PF00027; CNMP_binding; 1.		
DR	Pfam: PF00617; RASGEF; 1.		
DR	Pfam: PF00618; RASGEFN; 1.		
DR	PRINTS: PR00103; CAMPKINASE.		
DR	SMART; SM00100; CNMP; 1.		
DR	SMART; SM00147; RASGEF; 1.		
DR	SMART; SM00229; RASGEFN; 1.		
DR	PROSITE; PS50042; CNMP_BINDING_3; 1.		
DR	NON_TER		
FT	696		
FT	696		
SO	SEQUENCE		
	696 AA; 79440 MW; 4006240D4B85A747 CRC64;		
Query Match 48.6%; Score 855.5; DB 11; Length 696;			
Best Local Similarity 49.7%; Pred. No. 1.9e-69;			
Matches 168; Conservative 61; Mismatches 84; Indels 25; Gaps 3			
Qy	1	MAALAEQEGWTKGOVLVYVNSAGDAIGIQPDAKRGVATSGLNERLIVVNPQENHELIPR	60
Db	381	ISAVADKLGSGGGLLIIVKSNISGGEVYVLKSNVSVFTTLITIGRLPACRCRQDSLTPLR	440
Qy	61	DOLGPTVSAEELDIDVSAKDLGOLTDHDMSLFNSIHQVELIHVYLGPOHLRDVTANDE	120
Db	441	EOEGTGTGVCYGFELMSSKDIAVQMTYTWELFNCVHELILYHFG-RHNKRTKTANID	499
Db	121	FMRRRENQLQVWATLELCIPVGPRAQLLRFKFIKIAHLKQKLNISFFAVMFGLSNSA	180
Db	500	LFLRFNEIQEWWVVEVCLCSOLSKRVQLLKKFKIKIAACKEKYNLSFEAIVMGISNVA	559
Qy	181	ISRLAHMERLPHKYRKLXLSALERLDDSMNHRVRLAKLSPIVFPMLLKDMTFI	240
Db	560	VSRLALTEKRLPSFKFKYAEFESLMDSSRNHRATRLVAKLEPLIFPMLLIDMTFT	619

Oy		241	HGNCNTLVENLINFEXKMMMAAAMLLHCRCNHNVP.LSPLSRFXSHLEDQVARIStc	300
Dd		620	HEGNFTFDILVNFEKKMRRIATFARTVKRYSQ---PENPDAAAKNKHÖVR-----	669
Oy		301	SEOSLSTRPASTWAVVOOLKVINDQRELSRLSRELPE 338	
Dd		670	-----SYVOLANVIDNQRTLSQSHTLEP 693	
RESULT 12				
OQCWS52	ID			
OQCWS52	PRELIMINARY;	PRT:	699 AA.	
AC	OQCWS52;			
DT	01-JUN-2001 (TREMblrel .17, Created)			
DT	01-JUN-2001 (TREMblrel .17, Last sequence update)			
DT	01-MAR-2003 (TREMblrel .23, Last annotation update)			
DE	1300003D15RIk protein (Fragment).			
GN	CGEF2 OR 1300003D15RIK.			
OS	Mus musculus (Mouse).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
OX	NCB1-TaxID=10090;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=C57BL/6J; TISSUE=Liver;			
RX	MEDLINE=21085660; PubMed=11217651;			
RA	Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,			
RA	Aizawa T., Hara A., Fukunishi Y., Konno H., Adachi S., Fukuda S.,			
RA	Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaoka I.,			
RA	Saito T., Okazaki Y., Gojibori T., Bono H., Kasukawa T., Saito R.,			
RA	Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,			
RA	Fleischmann W., Gaasterland T., Gissi C., Kling B., Kochiya H.,			
RA	Kuehl P., Lewis S., Matsuo Y., Nikradlo I., Pesole G., Quackenbush J.,			
RA	Schmull L.M., Staudil F., Suzuki R., Tomita M., Wagner L., Washio T.,			
RA	Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,			
RA	Blake J., Boiffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,			
RA	Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,			
RA	Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,			
RA	L Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombarats P.,			
RA	Nardone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,			
RA	Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,			
RA	Suzuki H., Toyo-Oka K., Wang K.H., Weltz C., Whitlaker C., Wilming L.,			
RA	Wysshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaku S.,			
RA	Hayashizaki Y.;			
RT	*Functional annotation of a full-length mouse cDNA collection.*;			
RL	Nature 409:685-690(2001).			
DR	EMBL: AK004874; BMB23633.1; -.			
DR	MGD: MGI:1917723; CGE#2.			
DR	InterPro: IPR002373; CAMP_kin.			
DR	InterPro: IPR000595; cNMP_binding.			
DR	InterPro: IPR000651; RasGEF.			
DR	InterPro: IPR001895; RasGEF_CDC25.			
DR	pfam: PF00027; cNMP_binding; 1.			
DR	pfam: PF00617; RasGEF_1.			
DR	pfam: PF00618; RasGEFN; 1.			
DR	PRINTS: PR00103; CAMPkinASE.			
DR	SMART: SM00100; cNMP; 1.			
DR	SMART: SM00147; RasGEF; 1.			
DR	SMART: SM00229; RasGEFN; 1.			
DR	PROSITE: PS50042; cNMP_BINDING_3; 1.			
FT	NON_TER			
FT	SEQUENCE	699 AA: 79799 MW: AD2602AB9A41EA4 CRC64;		
Query Match		48.6%: Score 855.5; DB 11; Length 699;		
Best Local Similarity		49.7%: Pred. No. 1,9e+69;		
Matches	168; Conservative	61; Mismatches	84; Indels	25; Gaps
Oy		1	MALAAQEDDWTCGOVLVKNYSAGDAIGIOPDARGVATSLGLNERLFVNVDQEVNELIPHP 60	
Dd		384	ISAVADKIGSGGGGLITIKNNSGEKVKVLLSKNDVSPTTLINGRLTACPSCREOPDSLTLP 443	
Oy		61	DOLGPVGSAGEGLDVSAADLAGQLTDHMSLFSNSTHQVELIHVVYLGCPQHILRVYTANLE 120	

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using SW model

Run on: October 7, 2003, 06:57:32 ; Search time 24 Seconds

(without alignments)
662.293 Million cell updates/sec

Title: US-09-856-679-2

Perfect score: 1759
Sequence: 1 MAALAQEDGWTGCVLVKVN.....QLKVIDNQRELSRLSRELP 338

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database: SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	566	32.2	1234	1 YNKS_CAEEL	P34578 caenorhabd1
2	283	16.1	1077	1 C3G_HUMAN	Q13905 homo sapien
3	261	14.8	1571	1 C3G_DROME	O77086 drosophila
4	225	12.8	1095	1 CC25_SACKL	002342 saccharomyc
5	208.5	11.9	911	1 STE6_SCHPO	P26674 schizosacch
6	206	11.7	1244	1 GNRP_RAT	P28818 rattus norv
7	201.5	11.5	1262	1 GNRP_MOUSE	P27671 mus musculu
8	199	11.3	1275	1 GNRP_HUMAN	Q13972 homo sapien
9	196.5	11.2	1373	1 CC25_CANAL	P43069 candida alb
10	193	11.0	1339	1 SOS1_MOUSE	O62245 mus musculu
11	193	11.0	1333	1 SOS1_HUMAN	007889 homo sapien
12	188.5	10.7	1332	1 SOS2_HUMAN	O07890 homo sapien
13	181	10.3	1232	1 SC25_YEAST	P14771 saccharomyc
14	178	10.1	1589	1 CC25_YEAST	P04821 saccharomyc
15	178	10.1	1595	1 SOS_DROME	P26675 drosophila
16	177.5	10.1	1297	1 SOS2_MOUSE	002384 mus musculu
17	175.5	10.0	777	1 RGL2_HUMAN	O15211 homo sapien
18	173.5	9.9	778	1 RGL2_MOUSE	O61193 mus musculu
19	169.5	9.6	538	1 BUD5_YEAST	P25300 saccharomyc
20	157.5	9.0	1435	1 LTEL_YEAST	P07866 saccharomyc
21	153.5	8.7	852	1 GND5_MOUSE	O03385 mus musculu
22	150	8.5	914	1 GND5_HUMAN	O12967 homo sapien
23	148.5	8.4	768	1 RGL1_HUMAN	O9n216 homo sapien
24	143	8.1	768	1 RGL1_MOUSE	O60655 mus musculu
25	142.5	8.1	895	1 GND5_RAT	O03386 rattus norv
26	108	6.1	814	1 OPHL_HUMAN	O9una1 homo sapien
27	94	5.3	910	1 SC15_YEAST	P22224 saccharomyc
28	94	5.3	1844	1 POLR_TYVVA	P20128 turnip yell
29	93	5.3	452	1 ETV6_HUMAN	P41212 turnip yell
30	92.5	5.3	1844	1 POLR_TYVVA	P10358 turnip yell
31	92	5.2	594	1 TFC5_YEAST	P46678 saccharomyc
32	92	5.2	687	1 NPH1_MOUSE	O9qy53 mus musculu
33	91.5	5.2	1844	1 POLR_TYVWC	P28477 turnip yell

34	89.5	5.1	426	1 AGA2_ECOLI	P42903 escherichia
35	89.5	5.1	880	1 SYV_BACST	P11931 bacillus st
36	88.5	5.0	282	1 TAUD_ECOLI	P37610 escherichia
37	88.5	5.0	1321	1 IFX3_DICDI	O15818 dictyostell
38	88.5	5.0	2594	1 7LES_DROVI	P20806 drosophila
39	88	5.0	839	1 Y422_MYCPN	P75175 mycoplasma
40	88	5.0	1165	1 POL_GALV	P21414 gibbon ape
41	87.5	5.0	3321	1 PCN2_HUMAN	O95613 homo sapien
42	87.5	5.0	3674	1 SPCR_HUMAN	O9nc66 homo sapien
43	87	4.9	376	1 CRPF_RHOSH	P54906 rhodobacter
44	86	4.9	192	1 RR4B_CYACA	O22029 cyanidium c
45	86	4.9	306	1 Y910_TREPA	O83880 treponema p

ALIGNMENTS

```

RESULT 1
ID YNKS_CAEEL STANDARD: PRT: 1234 AA.
AC P34578:
DT 01-FEB-1994 (Rel. 28, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Hypothetical protein T20G5.5 in chromosome III.
GN T20G5.5.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC SPRAIN-Bristol NZ;
RA Berts M., Smith A.;
RL Submitted (MAR-1994) to the EMBL/GenBank/DBJ databases.
RN [2]
RP REVISIONS.
RA Durbin R.;
RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
CC - SIMILARITY: Contains 1 DEP domain.
CC - SIMILARITY: Contains 2 cyclic nucleotide-binding domains.
CC - SIMILARITY: Contains 1 Ras-GEF domain.
CC - SIMILARITY: Contains 1 N-terminal Ras-GEF domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
-----
EMBL: Z30423; CAAB3013.2; .
DR HSSP: P00515; ZBPk.
DR Wormpep: T20G5.5; CE23992.
DR InterPro: IPR002373; CAMP_kin.
DR InterPro: IPR000595; CNMP_binding.
DR InterPro: IPR000591; DEP.
DR InterPro: IPR000651; RasGEFN.
DR InterPro: IPR001895; RasGEF_CDC25.
DR Pfam: PF00027; CNMP_binding; 2.
DR Pfam: PF00610; DEP; 1.
DR Pfam: PF00617; RasGEF; 1.
DR Pfam: PF00618; RasGEFN; 1.
DR PRINTS: PR00103; CAMPKINASE.
DR SMART: SM00100; CNMP; 2.
DR SMART: SM00049; DEP; 1.
DR SMART: SM00147; RasGEF; 1.
DR SMART: SM00229; RasGEFN; 1.
DR PROSITE: PS00888; CNMP_BINDING_1; FALSE_NEG.
DR PROSITE: PS00889; CNMP_BINDING_2; FALSE_NEG.
DR PROSITE: PS50042; CNMP_BINDING_3; 2.
DR PROSITE: PS50186; DEP; 1.

```

```

DR PROSITE: PS00720; RASGEF_1.
DR PROSITE: PS50009; RASGEF_CAT; 1.
DR PROSITE: PS50212; RASGEF_NTER; 1.
KW Hypothetical protein; Repeat.
FT DOMAIN 206 336 CNMP-BINDING 1.
FT DOMAIN 430 512 DEP.
FT DOMAIN 571 688 CNMP-BINDING 2.
FT DOMAIN 712 850 N-TERMINAL RAS-GEF.
FT DOMAIN 991 1224 RAS-GEF.
SQ SEQUENCE 1234 AA; 141816 MW; 0E0C301CC424A58 CRC64;

Query Match 32.2%; Score 566; DB 1; Length 1234;
Best Local Similarity 38.3%; Pred. No. 2.1e-40;
Matches 124; Conservative 61; Mismatches 109; Indels 30; Gaps 4;

QY 16 LKVNAGDAIGLQDPARGVATSLGNERLFVNPQVEHLLPHPDGLPTVGSAG-LD 74
DB 929 LVEVKNGEKLEFSPDRALPTVLSKSLVYNNREIPLVMEQNGPFTSSSHSLH 988
QY 75 LVSAKDLAQQLDHDMSLFNSHQVELHYVGPQHLRVYTNALERFRRENELOVWA 134
DB 989 LIDSOELAHQLFLFHLQLRSTDSNELLYOVIGRESFPLSMFNLDLVRRFNEVOHWST 1048
QY 135 TELCLCPVGPGRROLRKFKILAAHLKEQKNLSFPAVMFGISNSAISRLAHTWELPRK 194
DB 1049 TELLLATEEN-RKEILLKFTISTATAREYRDLYVAITLGLSHTSISRLFTWSKLPPA 1107
QY 195 VRKLSALERLDPSNNHRYRLALAKLSPPVLPFMPLLDKMTFIEGNNHTLVENLNF 254
DB 1108 SLKTFELENLIDPTNNHMYRLVSKMSPTIPFVPLLLKDLMEHQNKSPYNLVNF 1167
QY 255 EKRMARARARMIHCRSHNPVPLSLRSRVSHLEDSDVARISSCSLSTRSPASTW 314
DB 1168 EKRMARAKIFRFRQKSO-----MDNGAEHEFIEPQ----- 1199
QY 315 AYYQOLKVIDNOERLSRELEP 338
DB 1200 SLIRNLVIDNOKKLMQSLYEIEP 1223

RESULT 2
C3G_HUMAN STANDARD; PRT; 1077 AA.
ID C3G_HUMAN
AC Q13905;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Guanine nucleotide-releasing factor 2 (C3g protein) (CRK SH3-binding
DE GNRP).
GN GNR2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM LONG), AND CHARACTERIZATION.
RC TISSUE=Placenta, and Spleen;
RX MEDLINE=94211880; PubMed=7512734;
RA Tanaka S., Morishita T., Hashimoto Y., Hattori S., Nakamura S.,
RA Shibuya M., Matuoka K., Takenawa T., Kurata T., Nagashima K.,
RA Matsuda M.;
RT "C3g, a guanine nucleotide-releasing protein expressed ubiquitously,
RT binds to the Src homology 3 domains of CRK and GRB2/ASH proteins.";
RT Proc. Natl. Acad. Sci. U.S.A. 91:3443-3447(1994).
RL [2]
RP SEQUENCE FROM N.A., FUNCTION, AND ALTERNATIVE SPLICING.
RX MEDLINE=95105157; PubMed=7806500;
RA Knudsen B., Feller S., Hanafusa H.;
RT "Four proline-rich sequences of the guanine-nucleotide exchange
RT factor C3g bind with unique specificity to the first Src homology 3
RT domain of Crk.";
RL J. Biol. Chem. 269:32781-32787(1994).
RN [3]

```

```

RP INTERACTION WITH CRK.
RX PubMed=8662907;
RA Matsuda M., Ota S., Tanimura R., Nakamura H., Matuoka K., Takenawa T.,
RA Nagashima K., Kurata T.;
RT "Interaction between the amino-terminal SH3 domain of CRK and its
RT natural target proteins.";
RL J. Biol. Chem. 271:14468-14472(1996).
CC -1- FUNCTION: Guanine nucleotide-releasing protein that binds to SH3
CC domain of CRK and GRB2/ASH. Transduces signals from CRK to
CC activate RAS.
CC -1- SUBUNIT: Interacts with CRK via its SH3-binding sites.
CC -1- ALTERNATIVE PRODUCTS:
CC Event-Alternative splicing; Named isoforms=2;
CC Name=Long;
CC IsoId=Q13905-1; Sequence=Displayed;
CC Name=Short;
CC IsoId=Q13905-2; Sequence=VSP_001822;
CC -1- TISSUE SPECIFICITY: Ubiquitously expressed in adult and fetus.
CC Expression is high in adult skeletal muscle and placenta and in
CC fetal brain and heart. Low levels of expression in adult and fetal
CC liver.
CC -1- SIMILARITY: Contains 1 Ras-GEF domain.
CC -1- SIMILARITY: Contains 1 N-terminal Ras-GEF domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sdb.ch/announce/
CC or send an email to license@isb-sdb.ch).
CC -----
DR EMBL: D21239; BAA04770.1; -.
DR Genbank: HGNC:4568; GNR2.
DR MIM: 600303; -.
DR GO: GO:0005515; P:protein binding activity; TAS.
DR GO: GO:0007169; P:transmembrane receptor protein tyrosine kin. . .; TAS.
DR InterPro: IPR000651; RASGEFN.
DR InterPro: IPR001895; RASGRF_CDC25.
DR Pfam: PF00617; RASGEF_1.
DR Pfam: PF00618; RASGEFN_1.
DR SMART: SM00147; RASGEFN_1.
DR SMART: SM00229; RASGEFN_1.
DR PROSITE: PS00720; RASGEF_1.
DR PROSITE: PS50009; RASGEF_CAT; 1.
DR PROSITE: PS50212; RASGEF_NTER; 1.
KW Guanine-nucleotide releasing factor; SH3-binding;
KW Alternative splicing.
FT DOMAIN 688 810
FT DOMAIN 840 1064
FT DOMAIN 963 966
FT SITE 281 292 SH3-BINDING.
FT SITE 451 462 SH3-BINDING.
FT SITE 538 549 SH3-BINDING.
FT SITE 606 617 SH3-BINDING.
FT VARSPPLIC 50 88
FT MISSING (in isoform Short).
FT /FTID=VSP_001822.
FT CONFLICT 137 137 P->R (IN REF. 2).
FT CONFLICT 183 183 E->G (IN REF. 2).
FT CONFLICT 217 217 S->C (IN REF. 2).
FT CONFLICT 254 254 T->S (IN REF. 2).
FT CONFLICT 287 287 P->T (IN REF. 1).
FT CONFLICT 355 355 D->G (IN REF. 2).
FT CONFLICT 636 636 G->V (IN REF. 2).
FT CONFLICT 698 698 D->N (IN REF. 2).
FT CONFLICT 846 847 EO->DE (IN REF. 2).
SQ SEQUENCE 1077 AA; 120604 MW; 7FF91348FC911CC CRC64;

Query Match 16.1%; Score 283; DB 1; Length 1077;
Best Local Similarity 28.3%; Pred. No. 3.4e-16;
Matches 79; Conservative 48; Mismatches 98; Indels 54; Gaps 8;
25 AIGLP-DARGVATSLGNERLFVNPQVEHLLPHPDGLPTVGSAGLDLVSAKDLAG 83

```

```

Db      818 ATSSOPLAARGVA-----ARPGTLHDFESH-----EIAE 846
Oy      84 QLTLDHDKSLFNSTHNOVELLYVGLPOHLRDTANTANLERFMRRNELQYVATELCLCPV 143
Db      847 QLTLLDALEF---YKTEIPVLLMAKQNEKSEKSNLOFHEFNFMNYSITIMIOEKA 903
Oy      144 GPAAQLLRKFIKTLAHLKEOKNLNSPFAVMFGLSNSAISRLAHTWELPKVRKLYSALE 203
Db      904 QDBERLLLRKFIKTLKHLKRLKLNNSYALITSAIDSAIRL--EMOKOTSEGLAETCT-- 959
Oy      204 RLDDPSMNRHVRALAKLSLPVLPMPPLLLKDMFTIHEGNHTLVENLINFEEK----- 256
Db      960 -LIDSSSFRAVRAALSEVEPPCIPYLGILITDTEFVHLNPNPYIDGKVNFSKRMQOFNI 1018
Oy      257 ---MRMARAAHMLHCHRSNHPVPLSRVSLHEDS 292
Db      1019 LDSMRCPFOA----HYDMRRNDITINFNDPSHLAEEA 1053

RESULT 3
C3G_DROME STANDARD: PRT: 1571 AA.
ID_C3G_DROME
AC 077086; Q9W3W3;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Guanine nucleotide-releasing factor 2 (CRK SH3-binding CNRP).
GN C3G OR CG3126.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM SHORT), AND CHARACTERIZATION.
RC TISSUE-Eye Imaginal disk;
RX MEDLINE=99094900; PubMed=9878058;
RA Ishihara S., Gaul U., Hanafusa H.;
RT "Activation of the Drosophila C3G leads to cell fate changes and
RT overproliferation during development, mediated by the RAS-MAPK
RT pathway and RAPI-1."
RL EMBO J. 18:145-155(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-Berkeley.
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Mortman J.R., Vandeil M.D., Zhang O., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazek R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Heit G., Nelson C.R., Miklos G.L.G.,
RA April J.F., Agbayan A., An H.-U., Andrews-Pinnock C., Baldwin D.,
RA Balow R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borokov D., Botchan M.R., Bouck J., Brokslein P., Brothier P.,
RA Burris K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cavley S., Dahke C., Davenport L.B., Davies P.,
RA de Pablo B., Deicher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
RA Foelel C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
RA Hostin N., Houston K.A., Howard T.J., Wei M.-H., Ibegwam C.,
RA Jallat M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Mekulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Paclob J.M.,

```

```

RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.C.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spreading A.C., Stapleton M., Strong R., Sun E.,
RA Svaydas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster."
RL Science 287:2185-2195(2000).
RN [3]
RP REVISIONS, AND ALTERNATIVE SPLICING.
RC STRAIN-Berkeley.
RX MEDLINE=22426069; PubMed=12537572;
RA Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,
RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochuk S.E.,
RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,
RA Bettencourt B.R., Celniker S.E., de Grey A.D.N.J., Drysdale R.A.,
RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.O.,
RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
RA Lewis S.E.;
RT "Annotation of the Drosophila melanogaster euchromatic genome: a
RT systematic review."
RL Genome Biol. 3:RESEARCH0083.1-RESEARCH0083.22(2002).
RN [4]
RP SEQUENCE OF 1036-1571 FROM N.A.
RC STRAIN-Berkeley; TISSUE-Ovary.
RX MEDLINE=22426066; PubMed=12537569;
RA Stapleton M., Carlson J.W., Brokslein P., Yu C., Champe M.,
RA George R.A., Guarini H., Kronmiller B., Paclob J.M., Park S., Wan K.H.,
RA Rubin G.M., Celniker S.E.;
RT "A Drosophila full-length cDNA resource."
RL Genome Biol. 3:RESEARCH0080.1-RESEARCH0080.8(2002).
CC -1- FUNCTION: Guanine nucleotide-releasing protein that binds to SH3
CC domain of CRK. Transduces signals from CRK to activate RAS. Also
CC involved in MAPK activation.
CC -1- ALTERNATIVE PRODUCTS:
CC Event-Alternative splicing; Named isoforms=2;
CC Name=Long;
CC IsoId=077086-1; Sequence=Displayed;
CC Name=Short;
CC IsoId=077086-2; Sequence=VSP_001823;
CC -1- TISSUE SPECIFICITY: Ubiquitous.
CC -1- DEVELOPMENTAL STAGE: Throughout development.
CC -1- SIMILARITY: Contains 1 Ras-GEF domain.
CC -1- SIMILARITY: Contains 1 N-terminal Ras-GEF domain.
CC -1- CAUTION: Ref.1 sequence differs from that shown due to a
CC frameshift in position 39.
CC -1- CAUTION: Ref.4 sequence differs from that shown due to a
CC frameshift in position 1366.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL: AF053358; AAC35280.1; ALT_FRAME.
DR EMBL: AE003438; AAF46200.2;
DR EMBL: AY113355; AAM9360.1; ALT_FRAME.
DR FLYBase: FBgn0026145; C3G.
DR GO: GO:0005088; F:RAS; guany[1]-nucleotide exchange factor activity; IMP.
DR GO: GO:0007265; P:RAS protein signal transduction; IMP.
DR InterPro: IPR000651; RASGEFN.
DR InterPro: IPR001895; RASGRF_C0C25.
DR Pfam: PF00617; RASGEF_1.
DR Pfam: PF00618; RASGEFN_1.
DR SMART: SM00147; RASGEF.
DR SMART: SM00229; RASGEFN_1.

```

DR PROSITE; PS00720; RASGEF; 1.
 DR PROSITE; PS50009; RASGEF_CAT; 1.
 DR PROSITE; PS50212; RASGEF_NTER; 1.
 KW Guanine-nucleotide releasing factor; SH3-binding;
 Developmental protein; Alternative splicing.
 FT DOMAIN 49
 FT DOMAIN 186
 FT DOMAIN 195
 FT DOMAIN 198
 FT DOMAIN 213
 FT DOMAIN 216
 FT DOMAIN 238
 FT DOMAIN 247
 FT DOMAIN 639
 FT DOMAIN 644
 FT DOMAIN 699
 FT DOMAIN 702
 FT DOMAIN 735
 FT DOMAIN 787
 FT DOMAIN 791
 FT DOMAIN 806
 FT DOMAIN 814
 FT DOMAIN 815
 FT DOMAIN 818
 FT DOMAIN 1170
 FT DOMAIN 1292
 FT DOMAIN 1462
 FT DOMAIN 1465
 FT SITE 546
 FT SITE 556
 FT SITE 820
 FT SITE 831
 FT SITE 924
 FT SITE 935
 FT SITE 986
 FT SITE 997
 FT VARSPLIC 668
 FT 718
 FT 19
 FT 93
 FT 154
 FT 202
 FT 202
 FT 243
 FT 243
 FT 450
 FT 453
 FT 477
 FT 477
 FT 807
 FT 818
 FT 851
 FT 1004
 FT 1092
 FT 1127
 FT 1127
 FT 1311
 FT 1311
 SQ SEQUENCE 1571 AA; 172049 MW; B9EF9E8670457684 CRC64;
 Query Match 14.8%; Score 261; DB 1; Length 1571;
 Best Local Similarity 33.3%; Pred. No. 4.5e-14;
 Matches 67; Conservative 43; Mismatches 73; Indels 18; Gaps 7;
 QY 64 GPTVGSAGC-----LDLVSAKDLAQLTDHDSLFNSIHQVELHYVLGPHLDVTTA 117
 DB 1321 GGCSGTAGGQNPISLDLKLST-ELAEQWTLDAELFT--KLEIEVLLFAKDCCEKSP 1376
 QY 118 NLERMRRENELQYVAVTELCIPVGPRAQLRKFEIKLAHLKQKNLNSFEAVMFGLS 177
 DB 1377 NLRKTEHFNRKMSYVARSKILRLDPAKERKHVNFYIKMKHLRMMNNYSTLLALSAD 1436
 QY 178 NSAISRLAHTEWR-LPHKVRKLYSALERLDPNMHRVRYRLALAKLSPVPIPFMDLLKD 236
 DB 1437 SGPIRL--EMQKGTTEVRSFCA-----LIDSSSFRAVYRQALATNPICPIYIGLLIOD 1490
 QY 237 MTFHEGNT--LVENLINFEEK 256
 DB 1491 LTFVHVGNDYLSKGVINFSK 1511
 RESULT 4
 CC25_SACKL STANDARD; PRT; 1095 AA.
 ID CC25_SACKL
 AC 002342;
 DT 01-JUL-1993 (Rel. 26, Last sequence update)
 DT 01-JUL-1993 (Rel. 26, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Cell division control protein 25 (Fragment).
 GN CDC25.
 OS Saccharomyces kluyveri (Yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.

OX NCBI_TaxID=4934;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=92354938; PubMed=1644315;
 RX Prigozy T., Gonzales E., Broek D.;
 RT "Identification and analysis of a DNA fragment from Saccharomyces
 kluyveri that can complement the loss of CDC25 function in
 Saccharomyces cerevisiae."
 RT
 RL Gene 117:67-72(1992).
 CC -! FUNCTION: PROMOTES THE EXCHANGE OF RAS-BOUND GDP BY GTP. THIS
 CC PROTEIN POSITIVELY CONTROLS THE LEVEL OF CELLULAR CAMP AT START,
 CC THE STAGE AT WHICH THE YEAST CELL DIVISION CYCLE IS TRIGGERED.
 CC -! SIMILARITY: Contains 1 Ras-GEF domain.
 CC -! SIMILARITY: Contains 1 N-terminal Ras-GEF domain.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC
 CC EMBL: M82964; AAA34479.1; -
 DR PIR; P01114; PC1114.
 DR InterPro; IPR000651; RASGEF.
 DR InterPro; IPR001895; RASGEF_CDC25.
 DR Pfam; PF00617; RASGEF; 1.
 DR Pfam; PF00618; RASGEF; 1.
 DR SMART; SM00147; RASGEF; 1.
 DR SMART; SM00229; RASGEF; 1.
 DR PROSITE; PS00720; RASGEF; 1.
 DR PROSITE; PS50009; RASGEF_CAT; 1.
 DR PROSITE; PS50212; RASGEF_NTER; 1.
 KW Guanine-nucleotide releasing factor; Cell division; Cell cycle;
 KW Mitosis; Transmembrane.
 FT NON_TER 1
 FT TRANSMEM 959
 FT DOMAIN 626
 FT DOMAIN 812
 FT SEQUENCE 1095 AA; 125605 MW; 7A8B2F90E31A44AC CRC64;
 Query Match 12.8%; Score 225; DB 1; Length 1095;
 Best Local Similarity 27.4%; Pred. No. 3.4e-11;
 Matches 75; Conservative 44; Mismatches 131; Indels 24; Gaps 7;
 QY 12 KGOVLYK---VNSADAIQLQPDARGAVATSLGLNRLTFVNPQEVNHLIPHPDOLGPTVG 68
 DB 752 KGRLSMKNLKNFVESINFSDGSSSTTV-----POSSRSVSAPVGVSSSTTG 800
 QY 69 -----SABGLDVSANDLAGQLTDHDSLFNSIHQVELHYVLGPHLDVTTANLERPMR 124
 DB 801 FRMRKLKLLD-IDSLDYAKQLTIKHSLEFYKISPECIDRTWGNKYNMGSKNTEFIS 859
 QY 125 RPNELQYVAVTELCIPVGPRAQLRKFEIKLAHLKQKNLNSFEAVMFGLSAISRL 184
 DB 860 NSNHLITNVSEFVAVQTIKRIQLIQFFINVAACHLNNPSSFTALISLYSPIYRL 919
 QY 185 AHTWERLPHKVRKLYSALERLDPNMHRVRYRLALAKLSP-PVPIPFMDLLKMTFHEG 243
 DB 920 KRTMAVAVDEYKKILKEELNLTMDSAKNFIRYRLKLSIGDFPCVFFGVYLSDLTFANG 979
 QY 244 NHTLVEN--LINF-EKRMAMARAARMLHHCRSH 273
 DB 980 NPDFLHNTVAVLNFGRVRIELEIKELISVYORSH 1013
 RESULT 5
 STE6_SCHPO STANDARD; PRT; 911 AA.
 ID STE6_SCHPO
 AC P26674; Q9JUM8;
 DT 01-AUG-1992 (Rel. 23, Last sequence update)
 DT 01-AUG-1992 (Rel. 23, Last sequence update)

DT 15-SEP-2003 (Rel. 42, last annotation update)
 DE Sleg protein.
 GN SLEG OR SPC1442.01 OR SPC1450.17.
 OS Schizosaccharomyces pombe (fission yeast)
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;
 OC Schizosaccharomyces.
 ON NCBI_TaxID=4896;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=90190870; PubMed=2107403;
 RA Hughes D.A., Fukui Y., Yamamoto M.;
 RT "Homologous activators of ras in fission and budding yeast.";
 RL Nature 344:355-357(1990).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX STRAIN=972;
 MEDLINE=21848401; PubMed=11859360;
 RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
 RA Sgouras J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
 RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
 RA Collins M., Connor R., Cronin A., Davis P., Felwell T., Fraser A.,
 RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
 RA Holtroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagsels K.,
 RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
 RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
 RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabbitts E.,
 RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
 RA Skellon J., Slimmons M., Squares R., Squares S., Stevens K.,
 RA Taylor R., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
 RA Woodward J., Volckaert G., Aert R., Robben J., Gymnopoulos B.,
 RA Wellens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
 RA Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H.,
 RA Borzym K., Langer I., Beck A., Lehnach H., Reinhardt R., Pohl T.M.,
 RA Eger P., Zimmermann W., Wedler H., Wambut R., Purnelle B.,
 RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,
 RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
 RA Lucas M., Rochet M., Gallard C., Tallada V.A., Garzon A., Thode G.,
 RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
 RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
 RA Cerretti L., Lowe T., McCormick W.R., Paulsen I., Potashkin J.,
 RA Siparova G.V., Ussery D., Barrell B.G., Nurse P.;
 RT "The genome sequence of Schizosaccharomyces pombe.";
 RL Nature 415:871-880(2002).
 CC -1- FUNCTION: PROMOTES THE EXCHANGE OF RAS-BOUND GDP BY GTP. THIS
 CC PROTEIN IS ESSENTIAL FOR MATING.
 CC -1- SIMILARITY: Contains 1 SH3 domain.
 CC -1- SIMILARITY: Contains 1 Ras-GEF domain.
 CC -1- SIMILARITY: Contains 1 N-terminal Ras-GEF domain.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: X53254; CAA37345.1; -
 CC EMBL: AL049559; CAB40184.1; -
 CC EMBL: AL031966; CAA21435.1; -
 CC PIR: S28098; S28098.
 CC GeneDB Spombe: SPC1442.01; -
 CC InterPro: IPR000651; RasGEF.
 CC InterPro: IPR001895; RasGEF_CDC25.
 CC InterPro: IPR001452; SH3.
 CC Pfam: PF00617; RasGEF_1.
 CC Pfam: PF00618; RasGEF_1.
 CC Pfam: PF00018; SH3_1.
 CC SMART: SM00147; RasGEF_1.
 CC SMART: SM00229; RasGEF_1.
 CC SMART: SM00326; SH3_1.
 CC PROSITE: PS00720; RasGEF_1.

DR PROSITE: PS50009; RasGEF_CAT; 1.
 DR PROSITE: PS50212; RasGEF_NTER; 1.
 DR PROSITE: PS50002; SH3_1.
 KW Guanine-nucleotide releasing factor; SH3 domain.
 FT DOMAIN 1 60
 FT DOMAIN 489 623 N-TERMINAL RAS-GEF.
 FT DOMAIN 663 894 RAS-GEF.
 SQ SEQUENCE 911 AA; 105185 MW; 5F8C12020C4B753F CRC64;
 Query Match 11.98; Score 208.5; DB 1; Length 911;
 Best Local Similarity 26.28; Pred. No. 6; Be-10;
 Matches 63; Conservative 43; Mismatches 107; Indels 27; Gaps 5;
 QY 51 QEVHELIPH--PQGLGTVASAGLDVSAKLAGOGLTDHDSVFNHGVLEHLYLGP 108
 Db 638 QLEEDLSLHNSPD---PIIKDELVLILPRETIKQCLIEFQSFHSIISIOPLTKIWD- 693
 QY 109 QHLRDVYTANLIERFMRR-----FNELOYVWATLCLCPVGPRAQLRKFKLAHL 160
 Db 694 -----NLNFSPEKSTFYLSNHLVNFYETETIVGEPPRRRTNVLAFYQVCDYL 744
 QY 161 KEQKRLNSFPAVNGLSNSAISRLAHTWELPKVKRLYSALERLDPSSNNHRYRLALA 220
 Db 745 RELNNFASLFSIISALNSSPIHRLKTMWMLNSKTLASFLLNNLTGARRKFNFSYRDCLE 804
 QY 221 KLSPPVFPMPILLKQWTFHEGNNHTLVENLNEPKRMMAARAARMLHGRS----HNPV 276
 Db 805 NCYLPVFPVGLVFTDTLFTLTKGNKFNQMNINEDKTKVTRILNETIKKFSQGYMENPI 864
 RESULT 6
 GNRP_RAT STANDARD; PRT; 1244 AA.
 ID GNRP_RAT
 AC P28818;
 DT 01-DEC-1992 (Rel. 24, Created)
 DT 01-DEC-1992 (Rel. 24, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Guanine nucleotide releasing protein (GNRP) (P140 Ras-GRF).
 GN RASGRPL
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 ON NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=92350260; PubMed=1379346;
 RA Shou C., Farusworth C.L., Neel B.G., Feig L.A.;
 RT "Molecular cloning of cDNAs encoding a guanine-nucleotide-releasing
 RT factor for Ras p21.";
 RL Nature 358:351-354(1992).
 CC -1- FUNCTION: PROMOTES THE EXCHANGE OF RAS-BOUND GDP BY GTP.
 CC -1- SIMILARITY: Contains 1 DBL-homology (DH) domain.
 CC -1- SIMILARITY: Contains 2 PH domains.
 CC -1- SIMILARITY: Contains 1 Ras-GEF domain.
 CC -1- SIMILARITY: Contains 1 IO domain.
 CC -1- SIMILARITY: Contains 1 N-terminal Ras-GEF domain.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: X67241; CAA47666.1; -
 CC PIR: S29083; S29083.
 CC InterPro: IPR001331; GDS_CDC24.
 CC InterPro: IPR000048; IQ_region.
 CC InterPro: IPR001849; PH.
 CC InterPro: IPR000651; RasGEF.
 CC InterPro: IPR001895; RasGEF_CDC25.

```
DR InterPro: IPR000219; RhoGEF.
DR Pfam: PF00612; IQ. 1.
DR Pfam: PF00169; PH. 2.
DR Pfam: PF00617; RasGEF. 1.
DR Pfam: PF00618; RasGEF. 1.
DR Pfam: PF00621; RhoGEF. 1.
DR SMART: SM00015; IQ. 1.
DR SMART: SM00233; PH. 2.
DR SMART: SM00147; RasGEF. 1.
DR SMART: SM00229; RasGEF. 2.
DR SMART: SM00325; RhoGEF. 1.
DR PROSITE: PS00741; DH_1. 1.
DR PROSITE: PS50010; DH_2. 1.
DR PROSITE: PS00720; RasGEF. 1.
DR PROSITE: PS50096; IQ. 1.
DR PROSITE: PS50003; PH_DOMAIN. 2.
DR PROSITE: PS50009; RasGEF_CAT. 1.
DR PROSITE: PS50212; RasGEF_NTER. 1.
KM Guanine-nucleotide releasing factor; Repeat.
FT DOMAIN 22 129
FT DOMAIN 204 229
FT DOMAIN 240 426
FT DOMAIN 456 582
FT DOMAIN 629 743
FT DOMAIN 1009 1241
SQ SEQUENCE 1244 AA; 142666 MW; 4B647879E842AF6B CRC64;

Query Match 11.7%; Score 206; DB 1; Length 1244;
Best Local Similarity 23.9%; Pred. No. 1.7e-09;
Matches 68; Conservative 56; Mismatches 144; Indels 16; Gaps 6;

QY 28 LQPDARVATSL--GLNRLVYVNPQEVHELPHPDQGLPVYGSAGGLDYSAKLAQL 85
DB 963 LIPQKRAAANMRLTQEEITENSMDELTLMEGV----KTEPEENHSAEIAQL 1017
QY 86 TDHWSLNSIHQVELIHVYVGPQHL--RDVTANLRPMKRFPELQYVATEICLQPV 142
DB 1018 TLDDHLVKSIPYEEF----FGCGMKADKNERTPIYIKTRHFPHNISLASELRDE 1073
QY 143 PGPRALRLKFKILAHLEKQNLNSFPAVMEGLNSAISRLAHTWERLPHKRVKLYSAL 202
DB 1074 VSARASTIEKVAVADICRCHLNVALEITSSINRSALFRLKKTWLKVSQKTSKLPDKL 1133
QY 203 ERLDPSNNHRYRLAKLSPVYIPMPYLKDKTFHHEGHTLVEN-LINFEKRMMA 261
DB 1134 QLVSSDGRFKMLRFTLCNCPVYLGMYLTDLAFLEEGPNYTEGLVNFSSMRMIS 1193
QY 262 RAARMLHCRSHNPVPLSLRSRVSHLHEDSOVARISTSCESQL 305
DB 1194 HIRREIROFO-QTYKIEPQRPVQYLVDEFVLDESLYEASL 1236

RESULT 7
GNRP_MOUSE STANDARD: PRT: 1262 AA.
AC P27671:
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Guanine nucleotide releasing protein (GNRP) (Ras-specific nucleotide
DE exchange factor CDC25) (CDC25Mm).
DE RASGRF1 OR CDC25 OR GRF1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euteria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_Taxid=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BAUB/c;
RX MEDLINE=93010996; PubMed=1396590;
RA Cen H., Lowy D.D.;
RT "Isolation of multiple mouse cDNAs with coding homology to
RT Saccharomyces cerevisiae CDC25. Identification of a region related to
```

```
RT Bcr, Vav, Dbl and CDC24.";
RN EMBO J. 11:4007-4015(1992).
RL [2]
RN SEQUENCE OF 791-1262 FROM N.A.
RP STRAIN=Swiss; TISSUE=Brain;
RC MEDLINE=92289680; PubMed=1376246;
RX Martegani E., Vanoni M., Zippel R., Cocchetti P., Brambilla R.,
RA Ferrari C., Scurni E.P., Alberghina L.,
RT "Cloning by functional complementation of a mouse cDNA encoding a
RT homologue of CDC25, a Saccharomyces cerevisiae Ras activator.";
RN EMBO J. 11:2151-2157(1992).
RP [3]
RN SEQUENCE OF 1031-1226 FROM N.A.
RX MEDLINE=92357779; PubMed=1379731;
RA Wei W., Mosteller R.D., Sanyal P., Gonzales E., McKinney D.,
RA Dasgupta C., Li P., Liu B.X., Broek D.;
RT "Identification of a mammalian gene structurally and functionally
RT related to the CDC25 gene of Saccharomyces cerevisiae.";
RN Proc. Natl. Acad. Sci. U.S.A. 89:7100-7104(1992).
RL - FUNCTION: PROMOTES THE EXCHANGE OF RAS-BOUND GDP BY GTP.
CC - TISSUE SPECIFICITY: BRAIN.
CC - SIMILARITY: Contains 2 PH domains.
CC - SIMILARITY: Contains 1 DBL-homology (DH) domain.
CC - SIMILARITY: Contains 1 Ras-GEF domain.
CC - SIMILARITY: Contains 1 IQ domain.
CC - SIMILARITY: Contains 1 N-terminal Ras-GEF domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: L20899; AAA02741.1; -.
DR EMBL: X59868; CAA42525.1; -.
DR PIR: S28407; S28407.
DR MGD: MGI:99694; Rasgrf1.
DR InterPro: IPR001331; GDS_CDC24.
DR InterPro: IPR000048; IQ_region.
DR InterPro: IPR001849; PH.
DR InterPro: IPR000651; RasGEFN.
DR InterPro: IPR001895; RasGEF_CDC25.
DR InterPro: IPR000219; RhoGEF.
DR Pfam: PF00612; IQ. 1.
DR Pfam: PF00169; PH. 2.
DR Pfam: PF00617; RasGEF. 1.
DR Pfam: PF00618; RasGEFN. 1.
DR Pfam: PF00621; RhoGEF. 1.
DR SMART: SM00015; IQ. 1.
DR SMART: SM00233; PH. 2.
DR SMART: SM00147; RasGEF. 1.
DR SMART: SM00229; RasGEFN. 2.
DR SMART: SM00325; RhoGEF. 1.
DR PROSITE: PS00741; DH_1. 1.
DR PROSITE: PS50010; DH_2. 1.
DR PROSITE: PS00720; RasGEF. 1.
DR PROSITE: PS50096; IQ. 1.
DR PROSITE: PS50003; PH_DOMAIN. 2.
DR PROSITE: PS50009; RasGEF_CAT. 1.
DR PROSITE: PS50212; RasGEF_NTER. 1.
KM Guanine-nucleotide releasing factor; Repeat.
FT DOMAIN 22 130
FT DOMAIN 208 233
FT DOMAIN 244 430
FT DOMAIN 460 588
FT DOMAIN 635 749
FT DOMAIN 1027 1259
FT DOMAIN 1033 1033
SQ SEQUENCE 1262 AA; 144101 MW; 38BE68F7C228DC8 CRC64;

Query Match 11.5%; Score 201.5; DB 1; Length 1262;
```

Best Local Similarity 25.1%; Pred. No. 4.2e-09;
Matches 58; Conservative 51; Mismatches 115; Indels 7; Gaps 4;

OY 78 AKDLAQGLTDHMSLFFNSIHQVELI--HYVLGPOHLADVTANLERMRBRENELOVAVAT 135
DB 1028 ALLEIAEDLTLLDLVFSIPSEEEFGGWMKAEEYER--TPYIMKTKTHFNHVSNTIAS 1084
OY 136 ELCLCPGPRADLLRKFKILAAHLKEOKNLSEFAVMFGLSNSAISRLAHTWERLPHKV 195
DB 1085 EIIRNEDISARASALEKVAADVADICRCLHNYNAVLETTSSINSALFRLLKTKMLKSKOT 1144
OY 196 RKLYSALERLLDPSWNRHYRLALAKISPPVIPPMLLKDMPTIEGHNHTLVN-LINF 254
DB 1145 KSLDLKLOKLVSDDGRFKNRESLRNCDPPCVPLGMYLTDLVFEETGTPNYTEDGLVNF 1204
OY 255 EKRRMARAARMLHHCHSHNPVPLSPRSVSHLHEDSOVARISTCEOSL 305
DB 1205 SKRMWISHIIRIROPQ-OTTYKIDPOPKVIOYLDESFMLDEESLYESSL 1254

RESULT 8
GNRP_HUMAN STANDARD; PRT: 1275 AA.

AC 013972;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Guanine nucleotide releasing protein (GNRP) (Ras-specific nucleotide exchange factor CDC25).
GN RASGRF1 OR CDC25.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95129875; PubMed-7828890;
RA Wei W., Das B., Park W., Brock D.;
RT "Cloning and analysis of human CDNA encoding a 140-kDa brain guanine nucleotide-exchange factor, Cdc25GEF, which regulates the function of Ras".
RT Ras".
RL Gene 151,279-284(1994).
CC -!- FUNCTION: PROMOTES THE EXCHANGE OF RAS-BOUND GDP BY GTP.
CC -!- SIMILARITY: Contains 2 PH domains.
CC -!- SIMILARITY: Contains 1 DBL-homology (DH) domain.
CC -!- SIMILARITY: Contains 1 Ras-GEF domain.
CC -!- SIMILARITY: Contains 1 IO domain.
CC -!- SIMILARITY: Contains 1 N-terminal Ras-GEF domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@sib-sib.ch).
CC -----
CC EMBL: L26584; AAA58417.1; -;
CC PIR: A38985; A38985.
CC Genew: HGNC:9875; RASGRF1.
CC GK: Q13972; -;
CC GO: GO:0005088; F:RAS guanyl-nucleotide exchange factor activity; TAS.
CC Interpro: IPR001331; GDS_CDC24.
CC Interpro: IPR000048; IO_region.
CC Interpro: IPR001849; PH.
CC Interpro: IPR000651; RASGEFN.
CC Interpro: IPR001895; RASGEF_CDC25.
CC Interpro: IPR000219; RHOGEF.
CC Pfam: PF00612; IO; 1.
CC Pfam: PF00169; PH; 2.
CC Pfam: PF00617; RASGEF_1.
CC Pfam: PF00618; RASGEFN_1.
CC Pfam: PF00621; RHOGEF_1.

DR SMART: SM00015; IO; 1.
DR SMART: SM00233; PH; 2.
DR SMART: SM00147; RASGEF_1.
DR SMART: SM00229; RASGEFN_2.
DR SMART: SM00325; RHOGEF_1.
DR PROSITE: PS00741; DH_1; 1.
DR PROSITE: PS00710; DH_2; 1.
DR PROSITE: PS00720; RASGEF_1.
DR PROSITE: PS50003; IO; 1.
DR PROSITE: PS50009; PH_DOMAIN; 2.
DR PROSITE: PS50009; RASGEF_CAT; 1.
DR PROSITE: PS50212; RASGEF_NTER; 1.
KW Guanine-nucleotide releasing factor; Repeat.
FT DOMAIN 22 129
FT DOMAIN 204 229
FT DOMAIN 240 426
FT DOMAIN 467 584
FT DOMAIN 644 762
FT DOMAIN 1040 1272
FT SEQUENCE 1275 AA; 145381 MW; 86C6F54A1E451F1 CRC64;

Query Match 11.3%; Score 199; DB 1; Length 1275;
Best Local Similarity 27.8%; Pred. No. 7.1e-09;
Matches 55; Conservative 41; Mismatches 98; Indels 4; Gaps 3;

OY 70 AEGIDLVSAKDAGQLTDHMSLFFNSIHQVELIHYVLPQHL-RDVTANLERMRBRENE 128
DB 1033 AEPENISALEIAQGLTLLDLVFKKIPYEEF--FGGWMKLEKNERTPYIMKTKTHFN 1090
OY 129 LQYVATELCLCPVGPRAOLLRKFKILAAHLKEOKNLSEFAVMFGLSNSAISRLAHTW 188
DB 1091 ISNLIAEIIIRNEDINRVAIEKVAADVADICRCLHNYNAVLETTSSMNSAISRLKKTW 1150
OY 189 ERLPHKRLKLYSALERLLDPSWNRHYRLALAKISPPVIPPMLLKDMPTIEGHNHTLV 248
DB 1151 LKYSKOTKALIDKLQKLVSSGFRKNLREALKNCDDPPCVPLGMYLTDLAETEGTPNYT 1210
OY 249 EN-LINEKRMARAAR 265
DB 1211 EDGLVNSKRMISHIR 1228

RESULT 9
CC25_CANAL STANDARD; PRT: 1333 AA.

ID CC25_CANAL
AC P43069;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Cell division control protein 25.
GN CDC25 OR CSC25.
OS Candida albicans (Yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; mitosporic Saccharomycetales; Candida.
OX NCBI_TaxID=5476;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=792-1;
RX MEDLINE=93238685; PubMed-8477693;
RA Goldberg D., Marbach I., Gross E., Levitzki A., Simchen G.;
RT "A Candida albicans homolog of CDC25 is functional in Saccharomyces cerevisiae".
RL Eur. J. Biochem. 213:195-204(1993).
CC -!- FUNCTION: PROMOTES THE EXCHANGE OF RAS-BOUND GDP BY GTP. THIS PROTEIN POSITIVELY CONTROLS THE LEVEL OF CELLULAR CAMP AT START, THE STAGE AT WHICH THE YEAST CELL DIVISION CYCLE IS TRIGGERED.
CC -!- SIMILARITY: Contains 1 SH3 domain.
CC -!- SIMILARITY: Contains 1 Ras-GEF domain.
CC -!- SIMILARITY: Contains 1 N-terminal Ras-GEF domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its

Query Match 11.0%; Score 193; DB 1; Length 1319;
 Best Local Similarity 27.9%; Pred. No. 2,4e-08;
 Matches 60; Conservative 46; Mismatches 83; Indels 26; Gaps 9;

QY 65 PIV-----GSNEGDLVSAA--DLAQLTDHDSLEPNSHOVELIHYLGQOHLDTV 115
 DB 759 PIVEMHISRGHETDLDLHPIETARQITLESIDLYRAVQSELSGWTRED-KEIN 817
 QY 116 TAVLEPFRNFNLOVWVATLELCVPVS--PRAOLRKFIKLAHLKEQKINSPFVAV 173
 DB 818 SPNLKMRITHTLTLMF--EKCTIVENLEERVAVSRITELQVFOELNENGYLEV 875
 QY 174 FGLSNSAISRLAHTWRLPRKVKLYSALERLLDPSNMHRYRLAKL--SPVPIPFM 230
 DB 876 SAMNSPVYKRLDTFQEDIPSRQKI--LEEAEHLSGDH--YKKYLAKLSINPCVPFF 930
 QY 231 PLLKDMTFIEGNGHTLV---ENLNEFKMRMA 261
 DB 931 GIVLTNLTKEGNEPVLRRHGKELINFSKRRVA 965

RESULT 11
 SOST_HUMAN STANDARD; PRT; 1333 AA.
 AC 007889;
 DT 15-JUL-1999 (Rel. 38, Created)
 DT 15-JUL-1999 (Rel. 38, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Son of sevenless protein homolog 1 (SOS-1).
 GN SOS1.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=93262494; PubMed=8493579;
 RA Chardin P., Camonla J.H., Gale N.W., van Aelst L., Wigler M.H.,
 RA Bar-Sagi D.;
 RT "Human Seg1: a guanine nucleotide exchange factor for Ras that binds
 RT to GRB2.";
 RL Science 260:1338-1343(1993).
 RN [2]
 RP STRUCTURE BY NMR OF 422-551.
 RX MEDLINE=98043737; PubMed=9374522;
 RA Zheng J., Chen R.H., Corbhan-Garcia S., Cahill S.M., Bar-Sagi D.,
 RA Cowburn D.;
 RT "The solution structure of the pleckstrin homology domain of human
 RT SOS1. A possible structural role for the sequential association of
 RT diffuse B cell lymphoma and pleckstrin homology domains.";
 RL J Biol. Chem. 272:30340-30344(1997).
 RN [3]
 RP X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS) OF 198-551.
 RX MEDLINE=99005193; PubMed=9790532;
 RA Solason S.M., Mmual A.S., Uy M., Bar-Sagi D., Kuriyan J.;
 RT "Crystal structure of the db1 and pleckstrin homology domains from
 RT the human Son of sevenless protein.";
 RL Cell 95:259-268(1998).
 CC -1- FUNCTION: PROMOTES THE EXCHANGE OF RAS-BOUND GDP BY GTP.
 CC -1- SUBUNIT: INTERACTS WITH GRB2.
 CC -1- SIMILARITY: Contains 1 db1-homology (DH) domain.
 CC -1- SIMILARITY: Contains 1 PH domain.
 CC -1- SIMILARITY: Contains 1 Ras-GEF domain.
 CC -1- SIMILARITY: Contains 1 N-terminal Ras-GEF domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC -----
 DR EMBL: L13857; AAA35913.1; -
 DR PDB: 1AME; 25-FEB-98.
 DR PDB; 1DBH; 23-DEC-98.
 DR Genew; HGNC:11187; SOS1.
 DR MIM; 182530; -
 DR GO: GO:0005088; F:RAS guanyl-nucleotide exchange factor activity; TAS.
 DR GO: GO:0005100; F:Rho GTPase activator activity; TAS.
 DR GO: GO:0005089; F:Rho guanyl-nucleotide exchange factor activity; TAS.
 DR GO: GO:0007265; P:RAS protein signal transduction; TAS.
 DR InterPro: IPR001331; GDS_CDC24.
 DR InterPro: IPR004822; Histone_core.
 DR InterPro: IPR001849; PH.
 DR InterPro: IPR000651; RasGEFN.
 DR InterPro: IPR001895; RasGEF_CDC25.
 DR InterPro: IPR000219; RhoGEF.
 DR Pfam; PF00169; PH; 1.
 DR Pfam; PF00617; RasGEF; 1.
 DR Pfam; PF00618; RasGEFN; 1.
 DR Pfam; PF00621; RhoGEF; 1.
 DR SMART; SM00233; PH; 1.
 DR SMART; SM00147; RasGEF; 1.
 DR SMART; SM00229; RasGEFN; 1.
 DR SMART; SM00325; RhoGEF; 1.
 DR PROSITE; PS00741; DH_1; FALSE_NEG.
 DR PROSITE; PS50010; DH_2; 1.
 DR PROSITE; PS50003; PH_DOMAIN; 1.
 DR PROSITE; PS50009; RasGEF_CAT; 1.
 DR PROSITE; PS50212; RasGEF_NTER; 1.
 DR KW Guanine-nucleotide releasing factor; 3D-structure.
 FT DOMAIN 200
 FT DOMAIN 444 548
 FT DOMAIN 597 741
 FT DOMAIN 780 1019
 FT DOMAIN 1258 1261
 FT HELIX 201 224
 FT TURN 225 227
 FT HELIX 228 232
 FT TURN 234 236
 FT HELIX 239 246
 FT TURN 247 248
 FT HELIX 249 268
 FT TURN 269 269
 FT TURN 272 273
 FT HELIX 280 288
 FT TURN 289 292
 FT HELIX 293 302
 FT TURN 303 303
 FT TURN 305 306
 FT HELIX 307 316
 FT TURN 318 319
 FT HELIX 320 327
 FT TURN 329 330
 FT HELIX 331 337
 FT TURN 338 338
 FT HELIX 339 342
 FT TURN 343 344
 FT HELIX 345 363
 FT TURN 367 379
 FT TURN 380 380
 FT HELIX 381 392
 FT HELIX 394 402
 FT TURN 403 403
 FT TURN 409 419
 FT HELIX 420 426
 FT TURN 427 428
 FT STRAND 429 430
 FT STRAND 437 439
 FT HELIX 440 440
 FT STRAND 444 452
 FT TURN 453 454
 FT STRAND 459 464

```

CC  STRAND  467  473
CC  TURN    475  476
CC  TURN    481  482
CC  STRAND  488  495
CC  STRAND  500  503
CC  STRAND  513  516
CC  TURN    519  520
CC  STRAND  524  527
CC  TURN    531  545
CC  HELIX   546  549
CC  SEQUENCE 1333 AA; 152463 MW; C6B99CCAL1A8DE45 CRC64;

Query Match 11.0%; Score 193; DB 1; Length 1333;
Best Local Similarity 27.9%; Pred. No. 2.5e-08;
Matches 60; Conservative 46; Mismatches 83; Indels 26; Gaps 9;

QY  PTV-----GSAEGLDVSAR--DLAQLTDHDMLSNSTHIOVLIHYVAGPQHLRYVT 115
DB  PTVEMHISRPCHIEFDLTLHPILARQLLLESDDLRAVOPSELVGSVWTKED-KEIN 817
QY  116 TANLEFRMRRNELQYVATELCPCVPG--PRAQLRKFKLAHLKEOKNLNSFFAVM 173
DB  818 SPNLKMIKRIHTNLTLMF--EKCIYETENLEERVAVSRITIELQVPELNNFNGVLHV 875
QY  174 FGLSASISRLAHTEPRPHKYRKYSALERLDPSSNHRVYRLALATL---SPVYIPFM 230
DB  876 SAMNSSPYRLDHFPEQIPSRQKI--LEEAHELSEDP--YKTYLAKRLSRINPCVPEF 930
QY  231 PLLKDMTFIHGHNHTLV----ENLINFEMKRMMA 261
DB  931 GYLTLNLTKEGNEPVLRKRGKELINFSKRKVA 965

RESULT 12
SOS2_HUMAN STANDARD; PRT; 1332 AA.
ID SOS2_HUMAN STANDARD; PRT; 1332 AA.
AC O07890; O15503;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DR 15-SEP-2003 (Rel. 42, Last annotation update)
DE Son of sevenless protein homolog 2 (SOS-2).
GN SOS2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RA MEDLINE=93262494; PubMed=8493579;
RA Chardin P.; Camonis J.H.; Gale N.W.; van Aelst L.; Wigler M.H.;
RA Bar-Sagi D.;
RT "Human Sos1: a guanine nucleotide exchange factor for Ras that binds
RT to GRB2."
RL Science 260:1338-1343(1993).
RN [2]
RP SEQUENCE OF 724-1296 FROM N.A.
RC TISSUE=Placenta;
RA Fath I.; Apiau F.; Dutrillaux B.; Tocque B.;
RL Submitted (MAR-1996) to the EMBL/Genbank/DBJ databases.
CC -1- FUNCTION: PROMOTES THE EXCHANGE OF RAS-BOUND GDP BY GTP (BY
CC SIMILARITY).
CC -1- SIMILARITY: Contains 1 DBL-homology (DH) domain.
CC -1- SIMILARITY: Contains 1 PH domain.
CC -1- SIMILARITY: Contains 1 Ras-GEF domain.
CC -1- SIMILARITY: Contains 1 N-terminal Ras-GEF domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/

```

```

CC  or send an email to license@isb-sib.ch).
CC  -----
CC  EMBL; L13858; AAA35914.1; -
DR  EMBL; L20686; AAA91852.1; -
DR  HSSP; Q62245; IPMS.
DR  Genew; HGNC:11188; SOS2.
DR  MIM; 601247; -
DR  GO; GO:0005085; F:guanylyl-nucleotide exchange factor activity; NAS.
DR  GO; GO:0007264; P:small GTPase mediated signal transduction; NAS.
DR  InterPro; IPR001331; GDS_CDC24.
DR  InterPro; IPR001849; PH.
DR  InterPro; IPR000651; RasGEFN.
DR  InterPro; IPR001895; RasGEF_CDC25.
DR  InterPro; IPR000219; RhogEF.
DR  Pfam; PF00169; PH; 1.
DR  Pfam; PF00617; RasGEF; 1.
DR  Pfam; PF00618; RasGEFN; 1.
DR  Pfam; PF00621; RhogEF; 1.
DR  SMART; SM00233; PH; 1.
DR  SMART; SM00147; RasGEF; 1.
DR  SMART; SM00229; RasGEFN; 1.
DR  SMART; SM00325; RhogEF; 1.
DR  PROSITE; PS00741; DH_1; FALSE_NEG.
DR  PROSITE; PS50010; DH_2; 1.
DR  PROSITE; PS00720; RasGEF; 1.
DR  PROSITE; PS50003; PH_DOMAIN; 1.
DR  PROSITE; PS30009; RasGEF_CAT; 1.
DR  PROSITE; PS50212; RasGEF_NTER; 1.
DR  KW  Guanine-nucleotide releasing factor.
FT  DOMAIN 198 388
FT  PH.
FT  N-TERMINAL RAS-GEF.
FT  DOMAIN 595 739
FT  RAS-GEF.
FT  DOMAIN 778 1017
FT  POLY-PRO.
FT  POLY-PRO.
FT  POLY-PRO.
FT  POLY-PRO.
FT  DOMAIN 1180 1183
FT  POLY-PRO.
FT  DOMAIN 1203 1208
FT  POLY-PRO.
FT  CONFLICT 778 778 D -> H (IN REF. 2).
FT  CONFLICT 861 861 Q -> R (IN REF. 2).
FT  CONFLICT 948 948 R -> K (IN REF. 2).
FT  CONFLICT 999 999 S -> C (IN REF. 2).
FT  CONFLICT 1032 1032 S -> E (IN REF. 2).
FT  CONFLICT 1042 1042 T -> A (IN REF. 2).
FT  CONFLICT 1112 1114 CGS -> LC (IN REF. 2).
FT  CONFLICT 1159 1159 A -> G (IN REF. 2).
FT  CONFLICT 1296 1296 Q -> E (IN REF. 2).
SQ  SEQUENCE 1332 AA; 153030 MW; EE4B9E52F865AC9 CRC64;

Query Match 10.7%; Score 188.5; DB 1; Length 1332;
Best Local Similarity 24.8%; Pred. No. 6e-08;
Matches 60; Conservative 53; Mismatches 102; Indels 27; Gaps 9;

QY  29 QDARGVATSTGLNRLFLVNPQEVHELIPHPDQLGPTVGSAGEGLDVSAR--DLAQLT 86
DB  740 QAOANGVSHNI-----TFESPPIEWHISKP-----GQETPDMLTLPIELARQLT 787
QY  87 DHDMSLFNSIHQVELIHVYLGPOHLADYTTANLEFRMRRNELQYVATELCPCVPG-- 144
DB  788 LLESPLKRVKQPSSELVGSVWTKED-KEINSPMLKMIKRIHTNLTLMF--EKCIYAENFE 844
QY  145 PRAQLRKFKIKLAHLKEOKNLNSFFAVMFGLSNSAISRLAHTEPRPHKYRKYSALER 204
DB  845 ERVAVALSRITIELQVFDLNNFNGVLEIVSAVNSVYRLDHTFPALQERKKRI--LDE 901
QY  205 LIDPSWNH-RVYRLALAKISPPVIFPMELKDMTFIHGHNHTLV----ENLINFEMKRM 259
DB  902 AVELSQDHKKYLVKLSINPCVPEFGIYLTINLTKEGNDLFRKRGKDLINFSKRK 961

RESULT 13
QY  260 MA 261
DB  962 VA 963

```

SC25_YEAST
ID SC25_YEAST STANDARD: PRT: 1252 AA.
AC P14771: Q12037: 012065;
DT 01-APR-1990 (Rel. 14, Created)
DT 15-SEP-2003 (Rel. 42, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Guanine nucleotide exchange factor SCD25.
GN SCD25 OR SCD25 OR YL016W/YL017W OR L1309/L1305.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetaceae.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-W303;
RX MEDLINE=91094833; PubMed=1986220;
RA Damak F., Boy-Marcotte E., Le Roscouet D., Gullbaud R., Jacquet M.;
RT "SCD25, a CDC25-like gene which contains a Ras-activating domain and
RL Mol. Cell. Biol. 11:202-212(1991).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-S288C / FY23;
RX MEDLINE=96405918; PubMed=8810043;
RA Miosga T., Zimmermann F.K.;
RT "Sequence analysis of the CEN12 region of Saccharomyces cerevisiae on
RT a 4.3.7 kb fragment of chromosome XII including an open reading frame
RT homologous to the human cystic fibrosis transmembrane conductance
RT regulator protein CFTR.";
RL Yeast 12:693-708(1996).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN-S288C;
RX MEDLINE=97197984; PubMed=9046100;
RA Purnelle B., Goffeau A.;
RT "The sequence of 32kb on the left arm of yeast chromosome XII reveals
RT six known genes, a new member of the seripauperins family and a new
RT ABS transporter homologous to the human multidrug resistance
RT protein.";
RL Yeast 13:183-188(1997).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN-S288C / AB972;
RX MEDLINE=97313267; PubMed=9169871;
RA Johnson M., Hillier L., Riles L., Albermann K., Andre B., Ansoorge W.,
RA Benes V., Brueckner M., Deltus H., Dubois E., Duesterhoeft A.,
RA Ertlan K.-D., Floeth M., Goffeau A., Hedling U., Heumann K.,
RA Heuss-Neitzel D., Hilbert H., Hilger F., Kleine K., Koeltter P.,
RA Louis E.J., Messenguy F., Mewes H.-W., Miosga T., Moestl D.,
RA Mueller-Auer S., Nentwich U., Obermaier B., Piravandi E., Pohl T.M.,
RA Portetle D., Purnelle B., Rechmann S., Rieger M., Rinke M., Rose M.,
RA Scharfe M., Scherens B., Scholler P., Schwager C., Schwarz S.,
RA Underwood A.P., Urrutia-L.A., Vandenbol M., Verhaaselt P.,
RA Vierendeels F., Voet M., Volckaert G., Voss H., Wambutt E., Wedler E.,
RA Wedler H., Zimmermann F.K., Zollner A., Hani J., Hobeisel J.D.;
RT "The nucleotide sequence of Saccharomyces cerevisiae chromosome XII.";
RL Nature 387:87-90(1997).
RN [5]
RP SEQUENCE OF 667-1252 FROM N.A.
RC STRAIN-OL136;
RX MEDLINE=89306677; PubMed=2545538;
RA Boy-Marcotte E., Damak F., Camonis J., Garreau H., Jacquet M.;
RT "The C-terminal part of a gene partially homologous to CDC 25 gene
RT suppresses the cdc25-5 mutation in Saccharomyces cerevisiae.";
RL Gene 77:21-30(1989).
RN [6]
RP FUNCTION;
RX MEDLINE=90260633; PubMed=2188363;
RA Crechet J.B., Poulet J., Mistou M.-Y., Parmegiani A., Camonis J.,
RA Boy-Marcotte E., Damak F., Jacquet M.;
RT "Enhancement of the GDP-GTP exchange of Ras proteins by the carboxyl-
RT terminal domain of SCD25.";
RL Science 248:866-868(1990).

RN [7]
RP FUNCTION;
RX MEDLINE=91156312; PubMed=2000228;
RA Rey I., Schweighoffer F., Barlat I., Camonis J., Boy-Marcotte E.,
RA Gullbaud R., Jacquet M., Tocque B.;
RT "The COOH-domain of the product of the Saccharomyces cerevisiae SCD25
RT gene elicits activation of p21-ras proteins in mammalian cells.";
RL Oncogene 6:347-349(1991).
RN [8]
RP FUNCTION;
RX PubMed=8730097;
RA Boy-Marcotte E., Ikonomi P., Jacquet M.;
RT "SCD25, a dispensable Ras guanine nucleotide exchange factor of
RT Saccharomyces cerevisiae differs from CDC25 by its regulation.";
RL Mol. Biol. Cell 7:529-539(1996).
CC -1- FUNCTION: PROMOTES THE EXCHANGE OF RAS-BOUND GDP BY GTP.
CC -1- MISCELLANEOUS: SUPPRESSES THE CDC25-5 MUTATION IN YEAST (RESTORES
CC CAMP LEVEL) AND HAS SIMILAR FUNCTIONS AS CDC25.
CC -1- MISCELLANEOUS: Strain S288C has a natural frameshift in position
CC 91 which disrupt the gene coding for this protein and produces two
CC ORFs.
CC -1- SIMILARITY: Contains 1 SH3 domain.
CC -1- SIMILARITY: Contains 1 Ras-GEF domain.
CC -1- SIMILARITY: Contains 1 N-terminal Ras-GEF domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M26647; AAA16565.1; -
DR EMBL: X91488; CAA62775.1; ALT_FRAME.
DR EMBL: X91488; CAA62774.1; ALT_FRAME.
DR EMBL: X97560; CAA66173.1; ALT_FRAME.
DR EMBL: X97560; CAA66161.1; ALT_FRAME.
DR EMBL: 273121; CAA97461.1; ALT_FRAME.
DR EMBL: 273122; CAA97462.1; ALT_FRAME.
DR SGD: S0003939; SDC25.
DR InterPro: IPR000651; RASGEFN.
DR InterPro: IPR001895; RASGRF_CDC25.
DR InterPro: IPR001452; SH3.
DR Pfam: PF006617; RASGEF.1.
DR Pfam: PF006618; RASGEF.1.
DR Pfam: PF00018; SH3.1.
DR SMART: SM00147; RASGEF.1.
DR SMART: SM00229; RASGEF.1.
DR SMART: SM00326; SH3.1.
DR PROSITE: PS00720; RASGEF.1.
DR PROSITE: PS50009; RASGEF_CAT.1.
DR PROSITE: PS50212; RASGEF_NTER.1.
DR PROSITE: PS50002; SH3.1.
KW Guanine-nucleotide releasing factor; Cell cycle; Cell division;
KW SH3 domain.
FT DOMAIN 26 97
FT DOMAIN 782 914
FT DOMAIN 952 1199
FT DOMAIN 74 78
FT DOMAIN 433 436
FT VARIANT 583 589
FT CONFILCT 78 78
FT CONFILCT 304 304
FT CONFILCT 377 379
FT CONFILCT 385 386
FT CONFILCT 417 418
FT CONFILCT 553 553
FT CONFILCT 557 557
FT CONFILCT 782 784
FT CONFILCT 809 809
FT CONFILCT 875 875
FT CONFILCT 915 915
SH3.
N-TERMINAL RAS-GEF.
RAS-GEF.
POLY-ASN.
POLY-ARG.
DVTVKFI -> V (IN STRAIN OL136).
N -> NN (IN REF. 1).
L -> S (IN REF. 1).
KRL -> IVF (IN REF. 1).
HD -> MT (IN REF. 1).
SC -> FI (IN REF. 1).
D -> G (IN REF. 1).
E -> K (IN REF. 1).
SNM -> PTY (IN REF. 1 AND 5).
D -> N (IN REF. 1 AND 5).
L -> P (IN REF. 1 AND 5).
H -> L (IN REF. 1 AND 5).


```

Db 1310 ANQLVLEHDLVLRITWFECLDRAMGCKYKCNMGSPNITKFIANANTLNFVSHITVKA 1369
QY 142 VPGPRAQLRKFKLAHLKEOKNLNSFAVMFGLNSAISRLAHTMERLPHVKRLYSA 201
Db 1370 DKTRSKLTQYFTVAQHCKELNNSMRAIVSALYSPYRLKTKTWDLVSTSKDLKN 1429
QY 202 LERLLDPSNMHRYRLALAKLSP-PIVPMPLLLKDMTEIHGNNHTLVE--NLINFEKM 257
Db 1430 LNNLMDSKRNFFVYRELRSVTDVACVPFGYVYLSDLTFTFVGNPDLNHNSTNINPSKR 1489
QY 258 RKMA 261
Db 1490 TKIA 1493

RESULT 15
SOS_DROME STANDARD; PRT; 1595 AA.
ID SOS_DROME
AC P26675;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Son of sevenless protein.
GN SOS.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephyroidae; Drosophilidae; Drosophila.
OX NCBI_TaxID:7227;
RN 11
RP SEQUENCE FROM N.A.
RC STRAIN-Oregon-R;
RX MEDLINE-92141820; Pubmed-1736363;
RA Bonfini L., Karlovich C.A., Dasgupta C., Banerjee U.;
RT "The Son of sevenless gene product: a putative activator of Ras.";
RL Science 253:603-606(1992).
RN 12
RP SEQUENCE FROM N.A.
RX MEDLINE-92034991; Pubmed-1934068;
RA Rasm M.A., Bonfanti D.D.L., Dodson G.S., Laverly T.R., Rubin G.M.;
RT "Ras and a putative guanine nucleotide exchange factor perform
RT crucial steps in signaling by the sevenless protein tyrosine
RT kinase.";
RL Cell 67:701-716(1991).
CC -1- FUNCTION: PROMOTES THE EXCHANGE OF RAS-BOUND GDP BY GTP. SOS IS
CC IMPLICATED IN NEURONAL DEVELOPMENT.
CC -1- SUBUNIT: MAY FORM A COMPLEX WITH SEVENLESS AND DRK.
CC -1- SIMILARITY: Contains 1 DBL-homology (DH) domain.
CC -1- SIMILARITY: Contains 1 Ras-GEF domain.
CC -1- SIMILARITY: Contains 1 N-terminal Ras-GEF domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: M83931; AAB04680.1; -
CC EMBL: M77501; AAA28904.1; -
CC PDB: 1AEZ; 18-MAY-99.
CC FLYBASE: FBgn0001965; Soc.
CC InterPro: IPR001331; GDS_CDC24.
CC InterPro: IPR004822; Histone_CORE.
CC InterPro: IPR002119; Histone_H2A.
CC InterPro: IPR001849; PH.
CC InterPro: IPR000651; RasGEFN.
CC InterPro: IPR001895; RasGRF_CDC25.
CC InterPro: IPR000219; RhoGEF.
CC Pfam: PF00169; PH; 1.
CC Pfam: PF00617; RasGEF; 1.

```

```

DR Pfam; PF00618; RasGEFN; 1.
DR Pfam; PF00621; RhoGEF; 1.
DR PRINTS; PR00620; HISTONEH2A.
DR SMART; SM00233; PH; 1.
DR SMART; SM00147; RasGEFN; 1.
DR SMART; SM00229; RasGEFN; 1.
DR SMART; SM00325; RhoGEF; 1.
DR PROSITE; PS00741; DH_1; FALSE_NEG.
DR PROSITE; PS50010; DH_2; 1.
DR PROSITE; PS00720; RasGEF; 1.
DR PROSITE; PS50003; PH_DOMAIN; 1.
DR PROSITE; PS50009; RasGEF_CAT; 1.
DR PROSITE; PS50212; RasGEF_NTER; 1.
KW Guanine-nucleotide releasing factor; Neurogenesis; 3D-structure.
FT DOMAIN 636 791 N-TERMINAL RAS-GEF.
FT DOMAIN 828 1065 RAS-GEF.
FT DOMAIN 15 51 GLY-RICH.
FT DOMAIN 247 433 DH.
FT DOMAIN 479 587 PH.
FT DOMAIN 1511 1516 GLN-RICH.
FT DOMAIN 1525 1541 HIS-RICH.
FT CONFLICT 232 243 TSCPVPCHPERS -> HILSPPLSLPAQR
FT CONFLICT (IN REF. 2).
FT CONFLICT V -> P (IN REF. 2).
SQ SEQUENCE 1595 AA; 177837 MW; 33AE31F0767A219F CRC64;

Query Match 10.1%; Score 178; DB 1; Length 1595;
Best Local Similarity 25.3%; Pred. No. 6.1e-07;
Matches 55; Conservative 47; Mismatches 95; Indels 20; Gaps 6;

QY 50 POEVHLLPHPDGLGPTVGSABGLDLYSANDLAGOLTDHDSLFNSIHQVELIHVYLGPO 109
Db 810 PIHHLISVPDEDTLLTLLHPLLE-----LARQLLLEFEMVKNVKKPSLVG---SPW 857
QY 110 HLED--VTANLEFRMRFRNELQYVATELCIDCPVQPRQQLRKFKLAHLKEOKNLN 167
Db 858 TKRDKVEKSPNLKIKMHTTNTVTRWIEKSTEAENVEERLAINQRALEVMVMLELNFN 917
QY 168 SFAVMFGLNSAISRLAHTMERLPHVKRLYSALERTLDPSMNH-RVYRLALAKLSPV 226
Db 918 GLIStVAAMGTASYVRLRMTFOGLPERYKRF---LECCRSLSDHLKKYERLRSLNPPC 974
QY 227 IPPMPLLLKDMTEIHGNNHTLVEN--LINEKRMMA 261
Db 975 VPFGRYLTNIIHLDEGNPDLANTLNTLINSKRKRYA 1011

```

Search completed: October 7, 2003, 07:13:42
Job time : 27 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 7, 2003, 07:13:18 : Search time 402 Seconds
(without alignments)
765.056 Million cell updates/sec

Title: US-09-856-679-2
Perfect score: 1759
Sequence: 1 MAALAQEDGWMTKGVLYKVN.....QLKVIDNQRELSRLSRELEP 338

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 5728757 seqs, 909918778 residues

Total number of hits satisfying chosen parameters: 5728757

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

Pending_Patents_AA_Main:*

1: /cgn2_6/ptodata/1/paa/US06_COMB.pep.*
2: /cgn2_6/ptodata/1/paa/US06_COMB.pep.*
3: /cgn2_6/ptodata/1/paa/US07_COMB.pep.*
4: /cgn2_6/ptodata/1/paa/US08_COMB.pep.*
5: /cgn2_6/ptodata/1/paa/US08_COMB.pep.*
6: /cgn2_6/ptodata/1/paa/US08_COMB.pep.*
7: /cgn2_6/ptodata/1/paa/US08_COMB.pep.*
8: /cgn2_6/ptodata/1/paa/US08_COMB.pep.*
9: /cgn2_6/ptodata/1/paa/US08_COMB.pep.*
10: /cgn2_6/ptodata/1/paa/US08_COMB.pep.*
11: /cgn2_6/ptodata/1/paa/US08_COMB.pep.*
12: /cgn2_6/ptodata/1/paa/US08_COMB.pep.*
13: /cgn2_6/ptodata/1/paa/US09_COMB.pep.*
14: /cgn2_6/ptodata/1/paa/US09_COMB.pep.*
15: /cgn2_6/ptodata/1/paa/US09_COMB.pep.*
16: /cgn2_6/ptodata/1/paa/US09_COMB.pep.*
17: /cgn2_6/ptodata/1/paa/US09_COMB.pep.*
18: /cgn2_6/ptodata/1/paa/US09_COMB.pep.*
19: /cgn2_6/ptodata/1/paa/US09_COMB.pep.*
20: /cgn2_6/ptodata/1/paa/US09_COMB.pep.*
21: /cgn2_6/ptodata/1/paa/US097A_COMB.pep.*
22: /cgn2_6/ptodata/1/paa/US097A_COMB.pep.*
23: /cgn2_6/ptodata/1/paa/US098_COMB.pep.*
24: /cgn2_6/ptodata/1/paa/US098_COMB.pep.*
25: /cgn2_6/ptodata/1/paa/US099A_COMB.pep.*
26: /cgn2_6/ptodata/1/paa/US099A_COMB.pep.*
27: /cgn2_6/ptodata/1/paa/US100_COMB.pep.*
28: /cgn2_6/ptodata/1/paa/US101_COMB.pep.*
29: /cgn2_6/ptodata/1/paa/US102_COMB.pep.*
30: /cgn2_6/ptodata/1/paa/US103_COMB.pep.*
31: /cgn2_6/ptodata/1/paa/US104_COMB.pep.*
32: /cgn2_6/ptodata/1/paa/US106_COMB.pep.*

Prod. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1759	100.0	338	23	US-09-856-679-2 Sequence 2, Appl1

2	1759	100.0	881	32	US-60-452-680-21605	Sequence 21605, A
3	1759	100.0	881	32	US-60-455-444-7029	Sequence 7029, Ap
4	1759	100.0	881	32	US-60-461-762-180	Sequence 460, Ap
5	1759	100.0	881	32	US-60-465-241-7029	Sequence 7029, Ap
6	1759	100.0	881	32	US-60-470-166-1217	Sequence 1217, Ap
7	1759	100.0	923	1	PCT-US02-36151-29	Sequence 29, Appl
8	1759	100.0	923	32	US-60-452-680-21606	Sequence 21606, A
9	1759	100.0	923	32	US-60-455-444-7030	Sequence 7030, Ap
10	1759	100.0	923	32	US-60-461-762-181	Sequence 461, Ap
11	1759	100.0	923	32	US-60-465-241-7030	Sequence 7030, Ap
12	1759	100.0	923	32	US-60-470-166-1218	Sequence 1218, Ap
13	1755	99.8	881	22	US-09-791-537-29766	Sequence 29766, A
14	1748	99.4	881	18	US-09-422-999-12	Sequence 12, Appl
15	1748	99.4	881	22	US-09-791-537-64901	Sequence 64901, A
16	1664	94.6	884	22	US-09-422-999-10	Sequence 10, Appl
17	1664	94.6	884	22	US-09-791-537-64922	Sequence 64922, A
18	1644.5	93.5	897	21	US-09-724-676-84651	Sequence 84651, A
19	1644.5	93.5	897	21	US-09-724-676-84651	Sequence 84651, A
20	1643	93.4	859	32	US-60-245-228-224	Sequence 224, Ap
21	1563	88.9	592	30	US-10-408-765A-1711	Sequence 1711, Ap
22	1563	88.9	592	32	US-60-389-987-1711	Sequence 1711, Ap
23	1563	88.9	592	32	US-60-412-418-1711	Sequence 1711, Ap
24	1408	80.0	270	24	US-09-911-826A-8	Sequence 8, Appl1
25	1013	57.6	235	22	US-09-760-446A-1834	Sequence 1834, Ap
26	1013	57.6	235	22	US-09-760-446A-1834	Sequence 1834, Ap
27	1013	57.6	235	28	US-10-206-664-1834	Sequence 1834, Ap
28	1013	57.6	235	28	US-10-216-436-726	Sequence 726, Ap
29	863.5	49.1	552	32	US-60-245-228-298	Sequence 298, Ap
30	863.5	49.1	552	32	US-60-258-272-124	Sequence 124, Ap
31	863.5	49.1	552	32	US-09-422-999-18	Sequence 18, Appl
32	863.5	49.1	1011	22	US-09-791-537-108781	Sequence 108781, A
33	863.5	49.1	1011	28	US-10-219-051B-8943	Sequence 8943, Ap
34	863.5	49.1	1011	28	US-10-219-051B-8943	Sequence 8943, Ap
35	855.5	48.6	993	22	US-09-791-537-118054	Sequence 70854, A
36	832.5	47.3	436	22	US-09-791-537-118054	Sequence 118054, A
37	832.5	47.3	436	28	US-10-219-051B-8943	Sequence 8943, Ap
38	832.5	47.3	436	28	US-10-219-051B-8943	Sequence 8943, Ap
39	786.5	44.7	291	18	US-09-422-999-16	Sequence 16, Appl
40	696	39.6	522	32	US-60-212-664-431	Sequence 431, Ap
41	696	39.6	957	20	US-09-614-150-6456	Sequence 6456, Ap
42	696	39.6	957	22	US-09-791-537-29071	Sequence 29071, A
43	696	39.6	957	32	US-60-167-217-5571	Sequence 6571, Ap
44	696	39.6	957	32	US-60-173-464-5273	Sequence 5273, Ap
45	696	39.6	957	32	US-60-191-637-6477	Sequence 6477, Ap

ALIGNMENTS

RESULT 1
US-09-856-679-2

Sequence 2, Application US/09856679

GENERAL INFORMATION:

APPLICANT: INCYTE PHARMACEUTICALS, INC.

APPLICANT: HILLMAN, Jennifer L.

APPLICANT: TANG, Y. Tom

APPLICANT: BANDMAN, Olga

APPLICANT: LAL, Preeti

APPLICANT: YUE, Henry

APPLICANT: LU, Dying Alina M.

APPLICANT: BAUGHN, Mariah R.

APPLICANT: YANG, Junming

APPLICANT: AZIMZAI, Valda

TITLE OF INVENTION: GTPASE ASSOCIATED PROTEINS

FILE REFERENCE: PF-0629 PCT

CURRENT APPLICATION NUMBER: US/09/856, 679

PRIOR FILING DATE: 2002-01-23

PRIOR APPLICATION NUMBER: 60/109, 592

PRIOR FILING DATE: 1998-11-23

PRIOR APPLICATION NUMBER: 60/118, 610

PRIOR FILING DATE: 1999-02-04

PRIOR APPLICATION NUMBER: 60/127, 990

PRIOR FILING DATE: 1999-04-06

```
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: PERL Program
; SEQ ID NO 2
; LENGTH: 338
; TYPE: prt
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc-feature
; OTHER INFORMATION: Incyte ID No: 1259937CD1
US-09-856-679-2
```

```
Query Match      100.0%; Score 1759; DB 23; Length 338;
Best Local Similarity 100.0%; Pred. No. 4.3e-173;
Matches 338; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 1 MAALAQEDGWTGQVLYVNSAGDAIGLPDARGVATSLGINERLFVNNPOEVHELIPHP 60
DB 1 MAALAQEDGWTGQVLYVNSAGDAIGLPDARGVATSLGINERLFVNNPOEVHELIPHP 60
OY 61 DOLGPTVSAGDLVSAKDLAAGQLTDHWSLFNSIHQVELIHVYLGPOHLRDVTTANLE 120
DB 61 DOLGPTVSAGDLVSAKDLAAGQLTDHWSLFNSIHQVELIHVYLGPOHLRDVTTANLE 120
OY 121 RFRMRFNELQYVWATELCLCPVPGRAQLRKFTKLAHLKEOKNLNSFFAVMGLSNSA 180
DB 121 RFRMRFNELQYVWATELCLCPVPGRAQLRKFTKLAHLKEOKNLNSFFAVMGLSNSA 180
OY 181 ISRLAHTWERLPHKVRKLYSALERLDPSSMHRVYRLALAKSPVPIPFMPLLLKDMTFI 240
DB 181 ISRLAHTWERLPHKVRKLYSALERLDPSSMHRVYRLALAKSPVPIPFMPLLLKDMTFI 240
OY 241 HEGNHTLVENLINFEKMMARARMLHCHSHNPVPLSPRSRVSHLHEDSOVARISTC 300
DB 241 HEGNHTLVENLINFEKMMARARMLHCHSHNPVPLSPRSRVSHLHEDSOVARISTC 300
OY 301 SEOSLSTRSPASTWAVYQOLKVIDNORELSRSLRELP 338
DB 301 SEOSLSTRSPASTWAVYQOLKVIDNORELSRSLRELP 338
```

RESULT 2

```
US-60-452-680-21605
; Sequence 21605, Application us/60452680
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; APPLICANT: GROPE, Andrew
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; FILE REFERENCE: C1001450
; CURRENT APPLICATION NUMBER: US/60/452,680
; CURRENT FILING DATE: 2003-03-07
; NUMBER OF SEQ ID NOS: 116213
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 21605
; LENGTH: 881
; TYPE: prt
; ORGANISM: Homo sapiens
US-60-452-680-21605
```

```
Query Match      100.0%; Score 1759; DB 32; Length 881;
Best Local Similarity 100.0%; Pred. No. 1.9e-172;
Matches 338; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 1 MAALAQEDGWTGQVLYVNSAGDAIGLPDARGVATSLGINERLFVNNPOEVHELIPHP 60
DB 544 MAALAQEDGWTGQVLYVNSAGDAIGLPDARGVATSLGINERLFVNNPOEVHELIPHP 603
OY 61 DOLGPTVSAGDLVSAKDLAAGQLTDHWSLFNSIHQVELIHVYLGPOHLRDVTTANLE 120
DB 604 DOLGPTVSAGDLVSAKDLAAGQLTDHWSLFNSIHQVELIHVYLGPOHLRDVTTANLE 663
OY 121 RFRMRFNELQYVWATELCLCPVPGRAQLRKFTKLAHLKEOKNLNSFFAVMGLSNSA 180
DB 121 RFRMRFNELQYVWATELCLCPVPGRAQLRKFTKLAHLKEOKNLNSFFAVMGLSNSA 180
```

```
DB 664 RFRMRFNELQYVWATELCLCPVPGRAQLRKFTKLAHLKEOKNLNSFFAVMGLSNSA 723
OY 181 ISRLAHTWERLPHKVRKLYSALERLDPSSMHRVYRLALAKSPVPIPFMPLLLKDMTFI 240
DB 724 ISRLAHTWERLPHKVRKLYSALERLDPSSMHRVYRLALAKSPVPIPFMPLLLKDMTFI 783
OY 241 HEGNHTLVENLINFEKMMARARMLHCHSHNPVPLSPRSRVSHLHEDSOVARISTC 300
DB 784 HEGNHTLVENLINFEKMMARARMLHCHSHNPVPLSPRSRVSHLHEDSOVARISTC 843
OY 301 SEOSLSTRSPASTWAVYQOLKVIDNORELSRSLRELP 338
DB 844 SEOSLSTRSPASTWAVYQOLKVIDNORELSRSLRELP 881
```

RESULT 3

```
US-60-455-444-7029
; Sequence 7029, Application us/60455444
; GENERAL INFORMATION:
; APPLICANT: BEGOVICH, Ann
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; FILE REFERENCE: C1001455
; CURRENT APPLICATION NUMBER: US/60/455,444
; CURRENT FILING DATE: 2003-03-18
; NUMBER OF SEQ ID NOS: 50986
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7029
; LENGTH: 881
; TYPE: prt
; ORGANISM: Homo sapiens
US-60-455-444-7029
```

```
Query Match      100.0%; Score 1759; DB 32; Length 881;
Best Local Similarity 100.0%; Pred. No. 1.9e-172;
Matches 338; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 1 MAALAQEDGWTGQVLYVNSAGDAIGLPDARGVATSLGINERLFVNNPOEVHELIPHP 60
DB 544 MAALAQEDGWTGQVLYVNSAGDAIGLPDARGVATSLGINERLFVNNPOEVHELIPHP 603
OY 61 DOLGPTVSAGDLVSAKDLAAGQLTDHWSLFNSIHQVELIHVYLGPOHLRDVTTANLE 120
DB 604 DOLGPTVSAGDLVSAKDLAAGQLTDHWSLFNSIHQVELIHVYLGPOHLRDVTTANLE 663
OY 121 RFRMRFNELQYVWATELCLCPVPGRAQLRKFTKLAHLKEOKNLNSFFAVMGLSNSA 180
DB 664 RFRMRFNELQYVWATELCLCPVPGRAQLRKFTKLAHLKEOKNLNSFFAVMGLSNSA 723
OY 181 ISRLAHTWERLPHKVRKLYSALERLDPSSMHRVYRLALAKSPVPIPFMPLLLKDMTFI 240
DB 724 ISRLAHTWERLPHKVRKLYSALERLDPSSMHRVYRLALAKSPVPIPFMPLLLKDMTFI 783
OY 241 HEGNHTLVENLINFEKMMARARMLHCHSHNPVPLSPRSRVSHLHEDSOVARISTC 300
DB 784 HEGNHTLVENLINFEKMMARARMLHCHSHNPVPLSPRSRVSHLHEDSOVARISTC 843
OY 301 SEOSLSTRSPASTWAVYQOLKVIDNORELSRSLRELP 338
DB 844 SEOSLSTRSPASTWAVYQOLKVIDNORELSRSLRELP 881
```

RESULT 4

```
US-60-461-762-480
; Sequence 480, Application us/60461762
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; APPLICANT: GROPE, Andrew
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; FILE REFERENCE: C1001459
; CURRENT APPLICATION NUMBER: US/60/461,762
```


;
; CURRENT FILING DATE: 2003-04-11
; NUMBER OF SEQ ID NOS: 30297
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 480
; LENGTH: 881
; TYPE: PRF
; ORGANISM: Homo sapiens
US-60-461-762-480

Query Match 100.0%; Score 1759; DB 32: Length 881;
Best Local Similarity 100.0%; Pred. No. 1.9e-172;
Matches 338; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MALAODEDWTGQGVVVKVNSAGDAIGLDPDARGVATSLGNRLFFVNPQEVNELIPHP 60
DB 544 MALAODEDWTGQGVVVKVNSAGDAIGLDPDARGVATSLGNRLFFVNPQEVNELIPHP 603
QY 61 DOLGPTVGSAGEGLDVSADLAGQLTDHMSLFNSIHVELIHYVLGPOHLRDVTYANLE 120
DB 604 DOLGPTVGSAGEGLDVSADLAGQLTDHMSLFNSIHVELIHYVLGPOHLRDVTYANLE 663
QY 121 RFRARRNELOYWATLCLCPVPGPRAQLLRKFIKLAHLKEQKLNSEFAVMEGLSNSA 180
DB 664 RFRARRNELOYWATLCLCPVPGPRAQLLRKFIKLAHLKEQKLNSEFAVMEGLSNSA 723
QY 181 ISRLATWTERLPHKVKLYSALERLDPSPMNRHVYRLALAKSPVYIPFMPILLKDMFTI 240
DB 724 ISRLATWTERLPHKVKLYSALERLDPSPMNRHVYRLALAKSPVYIPFMPILLKDMFTI 783
QY 241 HEGNHTLVENLINFEMRMARAARMLHHCRRSHNPVPLSPLSRSVSHLHEDSOVARISTC 300
DB 784 HEGNHTLVENLINFEMRMARAARMLHHCRRSHNPVPLSPLSRSVSHLHEDSOVARISTC 843
QY 301 SEOSLSTRSPASTWATVVOOLKVIDNORELSRSLRELEP 338
DB 844 SEOSLSTRSPASTWATVVOOLKVIDNORELSRSLRELEP 881

RESULT 5
US-60-465-241-7029

;
; Sequence 7029, Application US/60465241
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; APPLICANT: BEGOVICH, Ann
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: RHEUMATOID ARTHRITIS, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001466
; CURRENT APPLICATION NUMBER: US/60/465, 241
; CURRENT FILING DATE: 2003-04-23
; NUMBER OF SEQ ID NOS: 258418
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7029
; LENGTH: 881
; TYPE: PRF
; ORGANISM: Homo sapiens
US-60-465-241-7029

Query Match 100.0%; Score 1759; DB 32: Length 881;
Best Local Similarity 100.0%; Pred. No. 1.9e-172;
Matches 338; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MALAODEDWTGQGVVVKVNSAGDAIGLDPDARGVATSLGNRLFFVNPQEVNELIPHP 60
DB 544 MALAODEDWTGQGVVVKVNSAGDAIGLDPDARGVATSLGNRLFFVNPQEVNELIPHP 603
QY 61 DOLGPTVGSAGEGLDVSADLAGQLTDHMSLFNSIHVELIHYVLGPOHLRDVTYANLE 120
DB 604 DOLGPTVGSAGEGLDVSADLAGQLTDHMSLFNSIHVELIHYVLGPOHLRDVTYANLE 663
QY 121 RFRARRNELOYWATLCLCPVPGPRAQLLRKFIKLAHLKEQKLNSEFAVMEGLSNSA 180
DB 664 RFRARRNELOYWATLCLCPVPGPRAQLLRKFIKLAHLKEQKLNSEFAVMEGLSNSA 723

QY 181 ISRLATWTERLPHKVKLYSALERLDPSPMNRHVYRLALAKSPVYIPFMPILLKDMFTI 240
DB 724 ISRLATWTERLPHKVKLYSALERLDPSPMNRHVYRLALAKSPVYIPFMPILLKDMFTI 783
QY 241 HEGNHTLVENLINFEMRMARAARMLHHCRRSHNPVPLSPLSRSVSHLHEDSOVARISTC 300
DB 784 HEGNHTLVENLINFEMRMARAARMLHHCRRSHNPVPLSPLSRSVSHLHEDSOVARISTC 843
QY 301 SEOSLSTRSPASTWATVVOOLKVIDNORELSRSLRELEP 338
DB 844 SEOSLSTRSPASTWATVVOOLKVIDNORELSRSLRELEP 881

RESULT 6
US-60-470-166-1217

;
; Sequence 1217, Application US/60470166
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; APPLICANT: GRUPE, Andrew
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: ALZHEIMER'S DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001466
; CURRENT APPLICATION NUMBER: US/60/470, 166
; CURRENT FILING DATE: 2003-05-09
; NUMBER OF SEQ ID NOS: 68617
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1217
; LENGTH: 881
; TYPE: PRF
; ORGANISM: Homo sapiens
US-60-470-166-1217

Query Match 100.0%; Score 1759; DB 32: Length 881;
Best Local Similarity 100.0%; Pred. No. 1.9e-172;
Matches 338; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MALAODEDWTGQGVVVKVNSAGDAIGLDPDARGVATSLGNRLFFVNPQEVNELIPHP 60
DB 544 MALAODEDWTGQGVVVKVNSAGDAIGLDPDARGVATSLGNRLFFVNPQEVNELIPHP 603
QY 61 DOLGPTVGSAGEGLDVSADLAGQLTDHMSLFNSIHVELIHYVLGPOHLRDVTYANLE 120
DB 604 DOLGPTVGSAGEGLDVSADLAGQLTDHMSLFNSIHVELIHYVLGPOHLRDVTYANLE 663
QY 121 RFRARRNELOYWATLCLCPVPGPRAQLLRKFIKLAHLKEQKLNSEFAVMEGLSNSA 180
DB 664 RFRARRNELOYWATLCLCPVPGPRAQLLRKFIKLAHLKEQKLNSEFAVMEGLSNSA 723
QY 181 ISRLATWTERLPHKVKLYSALERLDPSPMNRHVYRLALAKSPVYIPFMPILLKDMFTI 240
DB 724 ISRLATWTERLPHKVKLYSALERLDPSPMNRHVYRLALAKSPVYIPFMPILLKDMFTI 783
QY 241 HEGNHTLVENLINFEMRMARAARMLHHCRRSHNPVPLSPLSRSVSHLHEDSOVARISTC 300
DB 784 HEGNHTLVENLINFEMRMARAARMLHHCRRSHNPVPLSPLSRSVSHLHEDSOVARISTC 843
QY 301 SEOSLSTRSPASTWATVVOOLKVIDNORELSRSLRELEP 338
DB 844 SEOSLSTRSPASTWATVVOOLKVIDNORELSRSLRELEP 881

RESULT 7
PCT-US02-36151-29

;
; Sequence 29, Application PCT/US0236151
; GENERAL INFORMATION:
; APPLICANT: INCYTE GENOMICS, INC.
; APPLICANT: KABLE, Amy E.; SMARNKAR, Anita
; APPLICANT: GORVAD, Ann E.; HAFALIA, April J.A.
; APPLICANT: DUGGAN, Brendan M.; WARREN, Bridget A.
; APPLICANT: EMERLING, Brooke E.; ISON, Craig H.
; APPLICANT: NGUYEN, Daniel B.; LINDQUIST, Erika A.
; APPLICANT: LEE, Ernestine A.; YUE, Henry
; APPLICANT: YUE, Huidin; FORSYTHE, Ian J.

```

: APPLICANT: RAMKUMAR, Jayalaxmi; GRIFFIN, Jennifer A.
: APPLICANT: Li, Joanna X.; MARQUIS, Joseph P.
: APPLICANT: GIERTZEN, Kimberly J.; BAUGHN, Mariah R.
: APPLICANT: BOROMSKY, Mark L.; YAO, Monique G.
: APPLICANT: CHAWLA, Nandier K.; LEHR-MASON, Patricia M.
: APPLICANT: LAL, Preeti G.; GURURAJAN, Rajagopal
: APPLICANT: KHARE, Reena; BARRA, Saiteev
: APPLICANT: BECHA, Shanya D.; LEE, Soo Yeun
: APPLICANT: TRAN, Uyen K.; ELLIOTT, Vicki S.
: APPLICANT: SPRAGUE, William W.; TANG, Y. Tom
: APPLICANT: ZEBARJADIAN, Yeganeh; JIANG, Xin
: APPLICANT: JACKSON, Alan A.; BHATTI, Umesh G.
: APPLICANT: BURRILL, John D.; LEE, Sally
: APPLICANT: BLAKE, Julie J.; HO, Anne
: TITLE OF INVENTION: INTRACELLULAR SIGNALING MOLECULES
: FILE REFERENCE: PF-1275 PCT
: CURRENT APPLICATION NUMBER: PCT/US02/36151
: PRIOR FILING DATE: 2002-11-07
: PRIOR APPLICATION NUMBER: US 60/344,472
: PRIOR FILING DATE: 2001-11-09
: PRIOR APPLICATION NUMBER: US 60/334,558
: PRIOR FILING DATE: 2001-11-30
: PRIOR APPLICATION NUMBER: US 60/340,296
: PRIOR FILING DATE: 2001-12-14
: PRIOR APPLICATION NUMBER: US 60/343,557
: PRIOR FILING DATE: 2001-12-21
: PRIOR APPLICATION NUMBER: US 60/350,420
: PRIOR FILING DATE: 2002-01-18
: PRIOR APPLICATION NUMBER: US 60/351,927
: PRIOR FILING DATE: 2002-01-25
: NUMBER OF SEQ ID NOS: 104
: SOFTWARE: PERL Program
: SEQ ID NO 29
: LENGTH: 923
: TYPE: PRT
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: misc_feature
: OTHER INFORMATION: Incyte ID NO: 1004901CD1
: PCT-US02-36151-29

Query Match      100.0%; Score 1759; DB 1; Length 923;
Best Local Similarity 100.0%; Pred. No. 2e-172;
Matches 338; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 MAALAOEDGWTGKGVLYKVNNSAGDAIGLQDPARGVATSLGNERLFVYVNPQEVHELIPHP 60
DB 586 MAALAOEDGWTGKGVLYKVNNSAGDAIGLQDPARGVATSLGNERLFVYVNPQEVHELIPHP 645
OY 61 DOLGPTVGSARGDLVSAKDIAGOLTDHDSLFNSIHQVELIHVYLGPOHLRDVTTANLE 120
DB 646 DOLGPTVGSARGDLVSAKDIAGOLTDHDSLFNSIHQVELIHVYLGPOHLRDVTTANLE 705
OY 121 RFRMRFNELQYVWATELCLCPVPGRAOLLRKFYKLAHLKEQKNLSFFAVMFGLSNSA 180
DB 706 RFRMRFNELQYVWATELCLCPVPGRAOLLRKFYKLAHLKEQKNLSFFAVMFGLSNSA 765
OY 181 ISRLAHTWERLPHKVRKLYSALERLDPNMHRVRYRLALAKLSPVPIPFMPLLKDMTFI 240
DB 766 ISRLAHTWERLPHKVRKLYSALERLDPNMHRVRYRLALAKLSPVPIPFMPLLKDMTFI 825
OY 241 HEGNHTLVENLINFCKMMAARAARMLHCRSHNPVPLSPRSRVSHLHEDSOVARISTC 300
DB 826 HEGNHTLVENLINFCKMMAARAARMLHCRSHNPVPLSPRSRVSHLHEDSOVARISTC 885
OY 301 SEQSLSTRSPASTWAVYQOLKVIDNORELSRSLRELEP 338
DB 886 SEQSLSTRSPASTWAVYQOLKVIDNORELSRSLRELEP 923

RESULT 8
US-60-452-680-21606
```

```

: Sequence 21606, Application US/60452680
: GENERAL INFORMATION:
: APPLICANT: CARGILL, Michele
: APPLICANT: GRUPE, Andrew
: TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
: TITLE OF INVENTION: ALZHEIMER'S DISEASE, METHODS OF DETECTION AND USES THEREOF
: FILE REFERENCE: CU001450
: CURRENT APPLICATION NUMBER: US/60/452,680
: CURRENT FILING DATE: 2003-03-07
: NUMBER OF SEQ ID NOS: 116213
: SOFTWARE: FASTSEQ for Windows Version 4.0
: SEQ ID NO 21606
: LENGTH: 923
: TYPE: PRT
: ORGANISM: Homo sapiens
: US-60-452-680-21606
```

```

Query Match      100.0%; Score 1759; DB 32; Length 923;
Best Local Similarity 100.0%; Pred. No. 2e-172;
Matches 338; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```

OY 1 MAALAOEDGWTGKGVLYKVNNSAGDAIGLQDPARGVATSLGNERLFVYVNPQEVHELIPHP 60
DB 586 MAALAOEDGWTGKGVLYKVNNSAGDAIGLQDPARGVATSLGNERLFVYVNPQEVHELIPHP 645
OY 61 DOLGPTVGSARGDLVSAKDIAGOLTDHDSLFNSIHQVELIHVYLGPOHLRDVTTANLE 120
DB 646 DOLGPTVGSARGDLVSAKDIAGOLTDHDSLFNSIHQVELIHVYLGPOHLRDVTTANLE 705
OY 121 RFRMRFNELQYVWATELCLCPVPGRAOLLRKFYKLAHLKEQKNLSFFAVMFGLSNSA 180
DB 706 RFRMRFNELQYVWATELCLCPVPGRAOLLRKFYKLAHLKEQKNLSFFAVMFGLSNSA 765
OY 181 ISRLAHTWERLPHKVRKLYSALERLDPNMHRVRYRLALAKLSPVPIPFMPLLKDMTFI 240
DB 766 ISRLAHTWERLPHKVRKLYSALERLDPNMHRVRYRLALAKLSPVPIPFMPLLKDMTFI 825
OY 241 HEGNHTLVENLINFCKMMAARAARMLHCRSHNPVPLSPRSRVSHLHEDSOVARISTC 300
DB 826 HEGNHTLVENLINFCKMMAARAARMLHCRSHNPVPLSPRSRVSHLHEDSOVARISTC 885
OY 301 SEQSLSTRSPASTWAVYQOLKVIDNORELSRSLRELEP 338
DB 886 SEQSLSTRSPASTWAVYQOLKVIDNORELSRSLRELEP 923

RESULT 9
US-60-454-7030
: Sequence 7030, Application US/60455444
: GENERAL INFORMATION:
: APPLICANT: CARGILL, Michele
: APPLICANT: BEGOVICH, Ann
: TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
: TITLE OF INVENTION: RHEUMATOID ARTHRITIS, METHODS OF DETECTION AND USES THEREOF
: FILE REFERENCE: CU001455
: CURRENT APPLICATION NUMBER: US/60/455,444
: CURRENT FILING DATE: 2003-03-18
: NUMBER OF SEQ ID NOS: 50986
: SOFTWARE: FASTSEQ for Windows Version 4.0
: SEQ ID NO 7030
: LENGTH: 923
: TYPE: PRT
: ORGANISM: Homo sapiens
: US-60-454-7030

Query Match      100.0%; Score 1759; DB 32; Length 923;
Best Local Similarity 100.0%; Pred. No. 2e-172;
Matches 338; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 MAALAOEDGWTGKGVLYKVNNSAGDAIGLQDPARGVATSLGNERLFVYVNPQEVHELIPHP 60
DB 586 MAALAOEDGWTGKGVLYKVNNSAGDAIGLQDPARGVATSLGNERLFVYVNPQEVHELIPHP 645
```

```
OY 61 DOLGPTVSAEGLDVSAKDLAQGLTDHMSLFNSIHQVELIHVYLGPOHLRDVTTANLE 120
    |||
Db 646 DOLGPTVSAEGLDVSAKDLAQGLTDHMSLFNSIHQVELIHVYLGPOHLRDVTTANLE 705
OY 121 RFRMRNELOYWATETCLCPVPGPRAQLLRKFTIKLAHLKEOKNLSFPAVMEGJSNSA 180
    |||
Db 706 RFRMRNELOYWATETCLCPVPGPRAQLLRKFTIKLAHLKEOKNLSFPAVMEGJSNSA 765
OY 181 ISRLAHTWERLPHKVRKLYSALERLDPSSNNHRYRLALAKLSPPVLPFMPLLKDMFTI 240
    |||
Db 766 ISRLAHTWERLPHKVRKLYSALERLDPSSNNHRYRLALAKLSPPVLPFMPLLKDMFTI 825
OY 241 HEGNHTLVENLINF EKMMARAARMLHCRSHNPVPLSPRSRVSHLHEDSOVARISTC 300
    |||
Db 826 HEGNHTLVENLINF EKMMARAARMLHCRSHNPVPLSPRSRVSHLHEDSOVARISTC 885
OY 301 SEOSLSTRSPASTWAVYQOLKVIDNOEELSRLSRELEP 338
    |||
Db 886 SEOSLSTRSPASTWAVYQOLKVIDNOEELSRLSRELEP 923
```

RESULT 10

```
; Sequence 481, Application US/60461762
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; APPLICANT: GARGILL, Andrew
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; FILE REFERENCE: C1001459
; CURRENT APPLICATION NUMBER: US/60/461,762
; CURRENT FILING DATE: 2003-04-11
; NUMBER OF SEQ ID NOS: 30297
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 481
; LENGTH: 923
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-461-762-481
```

```
Query Match 100.0%; Score 1759; DB 32; Length 923;
Best Local Similarity 100.0%; Pred. No. 2e-172;
Matches 338; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 1 MAALAOEDGWTGKGOVLKVSAGDAIGLOPDAGVATSLGNERLFVYNFOEYHELIPHP 60
    |||
Db 586 MAALAOEDGWTGKGOVLKVSAGDAIGLOPDAGVATSLGNERLFVYNFOEYHELIPHP 645
OY 61 DOLGPTVSAEGLDVSAKDLAQGLTDHMSLFNSIHQVELIHVYLGPOHLRDVTTANLE 120
    |||
Db 646 DOLGPTVSAEGLDVSAKDLAQGLTDHMSLFNSIHQVELIHVYLGPOHLRDVTTANLE 705
OY 121 RFRMRNELOYWATETCLCPVPGPRAQLLRKFTIKLAHLKEOKNLSFPAVMEGJSNSA 180
    |||
Db 706 RFRMRNELOYWATETCLCPVPGPRAQLLRKFTIKLAHLKEOKNLSFPAVMEGJSNSA 765
OY 181 ISRLAHTWERLPHKVRKLYSALERLDPSSNNHRYRLALAKLSPPVLPFMPLLKDMFTI 240
    |||
Db 766 ISRLAHTWERLPHKVRKLYSALERLDPSSNNHRYRLALAKLSPPVLPFMPLLKDMFTI 825
OY 241 HEGNHTLVENLINF EKMMARAARMLHCRSHNPVPLSPRSRVSHLHEDSOVARISTC 300
    |||
Db 826 HEGNHTLVENLINF EKMMARAARMLHCRSHNPVPLSPRSRVSHLHEDSOVARISTC 885
OY 301 SEOSLSTRSPASTWAVYQOLKVIDNOEELSRLSRELEP 338
    |||
Db 886 SEOSLSTRSPASTWAVYQOLKVIDNOEELSRLSRELEP 923
```

RESULT 11

```
; Sequence 7030, Application US/60465241
; GENERAL INFORMATION:
```

```
; APPLICANT: CARGILL, Michele
; APPLICANT: BEGOVICH, Ann
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; FILE REFERENCE: C1001468
; CURRENT APPLICATION NUMBER: US/60/465,241
; CURRENT FILING DATE: 2003-04-23
; NUMBER OF SEQ ID NOS: 258418
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7030
; LENGTH: 923
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-465-241-7030
```

```
Query Match 100.0%; Score 1759; DB 32; Length 923;
Best Local Similarity 100.0%; Pred. No. 2e-172;
Matches 338; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 1 MAALAOEDGWTGKGOVLKVSAGDAIGLOPDAGVATSLGNERLFVYNFOEYHELIPHP 60
    |||
Db 586 MAALAOEDGWTGKGOVLKVSAGDAIGLOPDAGVATSLGNERLFVYNFOEYHELIPHP 645
OY 61 DOLGPTVSAEGLDVSAKDLAQGLTDHMSLFNSIHQVELIHVYLGPOHLRDVTTANLE 120
    |||
Db 646 DOLGPTVSAEGLDVSAKDLAQGLTDHMSLFNSIHQVELIHVYLGPOHLRDVTTANLE 705
OY 121 RFRMRNELOYWATETCLCPVPGPRAQLLRKFTIKLAHLKEOKNLSFPAVMEGJSNSA 180
    |||
Db 706 RFRMRNELOYWATETCLCPVPGPRAQLLRKFTIKLAHLKEOKNLSFPAVMEGJSNSA 765
OY 181 ISRLAHTWERLPHKVRKLYSALERLDPSSNNHRYRLALAKLSPPVLPFMPLLKDMFTI 240
    |||
Db 766 ISRLAHTWERLPHKVRKLYSALERLDPSSNNHRYRLALAKLSPPVLPFMPLLKDMFTI 825
OY 241 HEGNHTLVENLINF EKMMARAARMLHCRSHNPVPLSPRSRVSHLHEDSOVARISTC 300
    |||
Db 826 HEGNHTLVENLINF EKMMARAARMLHCRSHNPVPLSPRSRVSHLHEDSOVARISTC 885
OY 301 SEOSLSTRSPASTWAVYQOLKVIDNOEELSRLSRELEP 338
    |||
Db 886 SEOSLSTRSPASTWAVYQOLKVIDNOEELSRLSRELEP 923
```

RESULT 12

```
; Sequence 1218, Application US/60470166
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; APPLICANT: GARGILL, Andrew
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; FILE REFERENCE: C1001464
; CURRENT APPLICATION NUMBER: US/60/470,166
; CURRENT FILING DATE: 2003-05-09
; NUMBER OF SEQ ID NOS: 68617
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1218
; LENGTH: 923
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-470-166-1218
```

```
Query Match 100.0%; Score 1759; DB 32; Length 923;
Best Local Similarity 100.0%; Pred. No. 2e-172;
Matches 338; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 1 MAALAOEDGWTGKGOVLKVSAGDAIGLOPDAGVATSLGNERLFVYNFOEYHELIPHP 60
    |||
Db 586 MAALAOEDGWTGKGOVLKVSAGDAIGLOPDAGVATSLGNERLFVYNFOEYHELIPHP 645
OY 61 DOLGPTVSAEGLDVSAKDLAQGLTDHMSLFNSIHQVELIHVYLGPOHLRDVTTANLE 120
    |||
```

```
Db      646 DOLGPTVGSAGELDLVSAKDLAAGQLTDHDMSLFNSIHQVELIHVYLGPQHLDVTTANLE 705
Qy      121 RMRMRRENLQYVWATELCLCPVGPRAQLLRKFILAAHLKEOKNLNSFFAVMFGLSNSA 180
Db      706 RMRMRRENLQYVWATELCLCPVGPRAQLLRKFILAAHLKEOKNLNSFFAVMFGLSNSA 765
Qy      181 ISRLAHTWERLPHKRYKRLYSALERLLDPSMNRHYRYRLALAKISPPVYIPFMPILLKDMFTI 240
Db      766 ISRLAHTWERLPHKRYKRLYSALERLLDPSMNRHYRYRLALAKISPPVYIPFMPILLKDMFTI 825
Qy      241 HEGNHTLVENLINFEMKRMMAARAARMLHHCSSHNPVPLSPKRSRVSHLHEDSOVARISTC 300
Db      826 HEGNHTLVENLINFEMKRMMAARAARMLHHCSSHNPVPLSPKRSRVSHLHEDSOVARISTC 885
Qy      301 SEQSLSRSPASTWAVYQOLKVIDNQRELSRLSRELEP 338
Db      886 SEQSLSRSPASTWAVYQOLKVIDNQRELSRLSRELEP 923

RESULT 13
US-09-791-537-29766
; Sequence 29766, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Bionomix, Inc.
; APPLICANT: Danzer, Derek
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMB
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; LENGTH: 881
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-791-537-29766

Query Match      99.8%; Score 1755; DB 22; Length 881;
Best Local Similarity 99.7%; Pred. No. 4.9e-172;
Matches 337; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      1 MAALAQEDGWTGGOVLVYVNSAGDAIGLOPDARGVATSLGNERLFFVYNPOEVHELIPHP 60
Db      544 MAALAQEDGWTGGOVLVYVNSAGDAIGLOPDARGVATSLGNERLFFVYNPOEVHELIPHP 603
Qy      61 DOLGPTVGSAGELDLVSAKDLAAGQLTDHDMSLFNSIHQVELIHVYLGPQHLDVTTANLE 120
Db      604 DOLGPTVGSAGELDLVSAKDLAAGQLTDHDMSLFNSIHQVELIHVYLGPQHLDVTTANLE 663
Qy      121 RMRMRRENLQYVWATELCLCPVGPRAQLLRKFILAAHLKEOKNLNSFFAVMFGLSNSA 180
Db      664 RMRMRRENLQYVWATELCLCPVGPRAQLLRKFILAAHLKEOKNLNSFFAVMFGLSNSA 723
Qy      181 ISRLAHTWERLPHKRYKRLYSALERLLDPSMNRHYRYRLALAKISPPVYIPFMPILLKDMFTI 240
Db      724 ISRLAHTWERLPHKRYKRLYSALERLLDPSMNRHYRYRLALAKISPPVYIPFMPILLKDMFTI 783
Qy      241 HEGNHTLVENLINFEMKRMMAARAARMLHHCSSHNPVPLSPKRSRVSHLHEDSOVARISTC 300
Db      784 HEGNHTLVENLINFEMKRMMAARAARMLHHCSSHNPVPLSPKRSRVSHLHEDSOVARISTC 843
Qy      301 SEQSLSRSPASTWAVYQOLKVIDNQRELSRLSRELEP 338
Db      844 SEQSLSRSPASTWAVYQOLKVIDNQRELSRLSRELEP 881

RESULT 14
US-09-422-999-12
; Sequence 12, Application US/09422999
; GENERAL INFORMATION:
; APPLICANT: Kawasaki, Hiroaki
```

```
; APPLICANT: Graybiel, Ann
; APPLICANT: Housman, David
; TITLE OF INVENTION: Genes Integrating Signal Transduction Pathways
; FILE REFERENCE: MIT-103
; CURRENT APPLICATION NUMBER: US/09/422,999
; CURRENT FILING DATE: 1999-10-22
; EARLIER APPLICATION NUMBER: US 60/105,507
; EARLIER FILING DATE: 1998-10-23
; EARLIER APPLICATION NUMBER: US 60/108,685
; EARLIER FILING DATE: 1998-11-16
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 12
; LENGTH: 881
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-422-999-12

Query Match      99.4%; Score 1748; DB 18; Length 881;
Best Local Similarity 99.1%; Pred. No. 2.6e-171;
Matches 335; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy      1 MAALAQEDGWTGGOVLVYVNSAGDAIGLOPDARGVATSLGNERLFFVYNPOEVHELIPHP 60
Db      544 MAALAQEDGWTGGOVLVYVNSAGDAIGLOPDARGVATSLGNERLFFVYNPOEVHELIPHP 603
Qy      61 DOLGPTVGSAGELDLVSAKDLAAGQLTDHDMSLFNSIHQVELIHVYLGPQHLDVTTANLE 120
Db      604 DOLGPTVGSAGELDLVSAKDLAAGQLTDHDMSLFNSIHQVELIHVYLGPQHLDVTTANLE 663
Qy      121 RMRMRRENLQYVWATELCLCPVGPRAQLLRKFILAAHLKEOKNLNSFFAVMFGLSNSA 180
Db      664 RMRMRRENLQYVWATELCLCPVGPRAQLLRKFILAAHLKEOKNLNSFFAVMFGLSNSP 723
Qy      181 ISRLAHTWERLPHKRYKRLYSALERLLDPSMNRHYRYRLALAKISPPVYIPFMPILLKDMFTI 240
Db      724 ISRLAHTWERLPHKRYKRLYSALERLLDPSMNRHYRYRLALAKISPPVYIPFMPILLKDMFTI 783
Qy      241 HEGNHTLVENLINFEMKRMMAARAARMLHHCSSHNPVPLSPKRSRVSHLHEDSOVARISTC 300
Db      784 HEGNHTLVENLINFEMKRMMAARAARMLHHCSSHNPVPLSPKRSRVSHLHEDSOVARISTC 843
Qy      301 SEQSLSRSPASTWAVYQOLKVIDNQRELSRLSRELEP 338
Db      844 SEQSLSRSPASTWAVYQOLKVIDNQRELSRLSRELEP 881

RESULT 15
US-09-791-537-64901
; Sequence 64901, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Bionomix, Inc.
; APPLICANT: Danzer, Derek
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY ME
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 64901
; LENGTH: 881
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-791-537-64901

Query Match      99.4%; Score 1748; DB 22; Length 881;
Best Local Similarity 99.1%; Pred. No. 2.6e-171;
Matches 335; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy      1 MAALAQEDGWTGGOVLVYVNSAGDAIGLOPDARGVATSLGNERLFFVYNPOEVHELIPHP 60
```

Db	544	MAALAOEDGWTGQVLVKVNSAGDAIGLPDAGVATSLGLNERLFPVNPQEVHELIPHP	603
QY	61	DOLGPTVGSABEGLDVSAKDLAQLTBDHMSLFNSIHQVELJHYVLGPOHLRDVTTANLE	120
Db	604	DOLGPTVGSABEGLDVSAKDLAQLTBDHMSLFNSIHQVELJHYVLGPOHLRDVTTANLE	663
QY	121	RFWRRFNELQYVWATELCLCPVGPRAQLLRKFIKLAHLKEQKNLNSFFAVMFGLSNSA	180
Db	664	RFWRRFNELQYVWATELCLCPVGPRAQLLRKFIKLAHLKEQKNVNSFFAVMFGLSNSP	723
QY	181	ISRLAHTWERLPHKVKKLYSALERLDPSPMNRHYRLALAKLSPVYIPFMPPLLLKDMTEI	240
Db	724	ISRLAHTWERLPHKVKKLYSALERLDPSPMNRHYRLALAKLSPVYIPFMPPLLLKDMTEI	783
QY	241	HEGNHTLVENLINFEXRMMARARMLHHCRRSHNPVPLSPLSRVSJLHEDSOVARISTC	300
Db	784	HEGNHTLVENLINFEXRMMARARMLHHCRRSHNPVPLSPLSRVSJLHEDSOVARISTC	843
QY	301	SEOSLSTRSPASTWAYVQOLKVIDNQRELSRLSRELEP	338
Db	844	SEOSLSTRSPASTWAYVQOLKVIDNQRELSRLSRELEP	881

Search completed: October 7, 2003, 07:23:37
Job time : 404 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 7, 2003, 07:15:33 : Search time 66 Seconds
(without alignments)
810.242 Million cell updates/sec

Title: US-09-856-679-2
Perfect score: 1759
Sequence: 1 MAALAEDEGWTGQVLVKVN.....QLKVIDNGRELRLSHELPE 338

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 587654 seqs, 158212981 residues
Total number of hits satisfying chosen parameters: 587654

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

Published_Applications_AA:*
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep:*
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep:*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep:*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep:*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/1/pubpaa/US10C_NEW_PUB.pep:*
17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*
18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1408	80.0	270	10 US-09-911-826A-8	Sequence 8, Appl1
2	419.5	23.8	246	10 US-09-940-836A-6	Sequence 6, Appl1
3	418.5	23.8	238	10 US-09-922-1199A-9	Sequence 9, Appl1
4	409	23.3	1499	10 US-09-911-826A-2	Sequence 2, Appl1
5	405.5	23.1	261	9 US-09-822-635-4	Sequence 4, Appl1
6	405.5	23.1	261	16 US-10-176-306-32	Sequence 32, Appl1
7	396.5	22.5	307	10 US-09-911-826A-7	Sequence 7, Appl1
8	246.5	14.0	227	10 US-09-922-1199A-5	Sequence 5, Appl1
9	246.5	14.0	227	16 US-10-176-306-31	Sequence 31, Appl1
10	238	13.5	472	10 US-09-940-836A-2	Sequence 2, Appl1
11	237	13.5	473	15 US-10-275-173-2	Sequence 2, Appl1
12	229	12.6	59	9 US-09-864-761-39774	Sequence 39774, A
13	221	12.6	591	12 US-09-805-455-7	Sequence 7, Appl1
14	207.5	11.8	782	15 US-10-177-293-399	Sequence 399, App
15	204	11.6	244	10 US-09-911-826A-9	Sequence 9, Appl1

16	196	11.1	243	10 US-09-911-826A-11	Sequence 11, Appl1
17	194	11.0	709	14 US-10-118-328-4	Sequence 4, Appl1
18	193	11.0	684	10 US-09-765-298A-18	Sequence 18, Appl1
19	191.5	10.9	201	15 US-10-103-313-420	Sequence 420, App
20	190.5	10.8	581	12 US-09-805-455-8	Sequence 8, App
21	190	10.8	428	15 US-10-157-031-139	Sequence 139, App
22	186	10.6	164	15 US-10-103-313-543	Sequence 543, App
23	182	10.3	699	15 US-10-233-131-34	Sequence 34, Appl1
24	182	10.3	710	14 US-10-118-328-2	Sequence 2, Appl1
25	182	10.3	710	15 US-10-060-990-3	Sequence 3, Appl1
26	182	10.3	715	15 US-10-233-131-33	Sequence 33, Appl1
27	178	10.1	1589	10 US-09-801-368-64	Sequence 64, Appl1
28	177.5	10.1	249	10 US-09-911-826A-10	Sequence 10, Appl1
29	175.5	10.0	777	15 US-10-205-823-345	Sequence 345, Appl1
30	170	9.7	169	10 US-09-940-836A-4	Sequence 4, Appl1
31	169.5	9.6	608	12 US-09-805-455-6	Sequence 6, Appl1
32	166.5	9.5	609	12 US-09-805-455-4	Sequence 4, Appl1
33	166.5	9.5	615	12 US-09-805-455-2	Sequence 2, Appl1
34	166.5	9.5	664	12 US-09-805-455-5	Sequence 5, Appl1
35	166	9.4	528	9 US-09-925-297-806	Sequence 806, App
36	162.5	9.2	926	14 US-10-108-605-143	Sequence 143, App
37	161.5	9.2	673	16 US-10-176-306-2	Sequence 2, Appl1
38	157.5	9.0	142	10 US-09-764-868-1023	Sequence 1023, App
39	157.5	9.0	142	15 US-10-103-313-429	Sequence 429, App
40	155	8.8	158	15 US-10-103-313-404	Sequence 404, App
41	150	8.5	677	15 US-10-103-313-384	Sequence 384, App
42	150	8.5	902	10 US-09-922-1199A-2	Sequence 2, Appl1
43	148.5	8.4	725	14 US-10-118-328-5	Sequence 5, Appl1
44	143	8.1	141	15 US-10-275-173-4	Sequence 4, Appl1
45	132.5	7.5	60	9 US-09-864-761-44037	Sequence 44037, A

ALIGNMENTS

RESULT 1	US-09-911-826A-8				
Sequence 8, Appl1	Application US/09911826A				
Patent No. US20020143164A1					
GENERAL INFORMATION:					
APPLICANT: Roclin, Daniela and Pham, Nam					
TITLE OF INVENTION: RAS Activator Nucleic Acid Molecules, Polypeptides and					
FILE REFERENCE: DDW-5001-US					
CURRENT APPLICATION NUMBER: US/09/911, 826A					
CURRENT FILING DATE: 2002-02-26					
PRIOR APPLICATION NUMBER: PCT/CA00/00042					
PRIOR FILING DATE: 2000-01-20					
PRIOR APPLICATION NUMBER: 2,259, 830					
PRIOR FILING DATE: 1999-01-20					
NUMBER OF SEQ ID NOS: 27					
SOFTWARE: Patentin Ver. 2.1					
SEQ ID NO 8					
LENGTH: 270					
TYPE: PRT					
ORGANISM: Homo sapiens					
US-09-911-826A-8					
Query Match	80.0%	Score 1408	DB 10	Length 270	
Best Local Similarity	100.0%	Pred. No. 1.1e-138			
Matches 270	Conservative 0	Mismatches 0	Indels 0	Gaps 0	
QY	69	SAGGLDLSAKDLAAGLTDDHDSLFNSIHQVELIHVYLGPOHLDVTTANLSPFRFRNE	128		
DB	1	SAGGLDLSAKDLAAGLTDDHDSLFNSIHQVELIHVYLGPOHLDVTTANLSPFRFRNE	60		
QY	129	LQYVATLCLCPVGPRAQLLKRFLKLAHLKEOKNLNSFFAVMGSLNSAISRLAHTW	188		
DB	61	LQYVATLCLCPVGPRAQLLKRFLKLAHLKEOKNLNSFFAVMGSLNSAISRLAHTW	120		
QY	189	ERLPHKRYKLYSLERLDDSMNRYRYRLAKLSPVYFPMLLKDMFTIHGNTLV	248		
DB	121	ERLPHKRYKLYSLERLDDSMNRYRYRLAKLSPVYFPMLLKDMFTIHGNTLV	180		

QY	249	ENLNEEKMMAAARAAMHHCRRNPVLSPLRSVSHLHDSQYARISTGEOSLSTR	308
QY	181	ENLNEEKMMAAARAAMHHCRRNPVLSPLRSVSHLHDSQYARISTGEOSLSTR	240
DB	309	SPASTMAYVOOLKVIDNORELSRLSRELEP	338
QY	241	SPASTMAYVOOLKVIDNORELSRLSRELEP	270
DB			

RESULT 2
US-09-940-836A-6

? Sequence 6, Application US/09940836A
 ? Patent No. US200201466800A1
 ? GENERAL INFORMATION:
 ? APPLICANT: Millennium Pharmaceuticals, Inc.
 ? APPLICANT: Cutlis, Roy
 ? TITLE OF INVENTION: 48921, A NOVEL HUMAN GTP RELEASING
 ? TITLE OF INVENTION: FACTOR AND USES THEREFOR
 ? FILE REFERENCE: 38155-20031.00
 ? CURRENT APPLICATION NUMBER: US/09/940,836A
 ? CURRENT FILING DATE: 2001-08-27
 ? PRIOR APPLICATION NUMBER: US 60/228,760
 ? PRIOR FILING DATE: 2000-08-30
 ? NUMBER OF SEQ ID NOS: 12
 ? SOFTWARE: FastSeq for Windows Version 4.0
 ? SEQ ID NO 6
 ? LENGTH: 246
 ? TYPE: PRT
 ? ORGANISM: Artificial Sequence
 ? FEATURE:
 ? OTHER INFORMATION: Consensus amino acid sequence
 ? US-09-940-836A-6

Query Match	23.88;	Score 419.5;	DB 10;	Length 246;
Best Local Similarly	43.38;	Pred. No. 2.2e-35;		
Matches 91; Conservative	40;	Mismatches 58;	Indels 21;	Gaps 5;

```

QY      73 LDD--VSADLAGOLTDDHMSLFNSHQVELIHY-----LGPOLHLDVYTANLEREMR 124
Db      8 LDIHLDPPELAEOLTLDDLELKKIEPSECLGYAWSNREKKGKEN---SPNLEKFIQ 63

```

0y 125 RENEQYAWATELCCPVBPRAOLRRFKFLAAHKEOKUNLSFEAVMFSGNSASIRL 184
|||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Ddb 64 RENNISYWATEIISSEKPOORAKITIEKFLVAQHCRELNPNFSLMAIVGSSNSSIRYL 123

QY 185 AHTEERLPNHRVVKRYSALAELEEDPWNHRYRLAKLS-----PVIFEMPLILKDM 237
|||:::|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 124 KKTWEKVRETKTLEEELSELMDPSNNKYNTRELKSCGTGVNHSDPPCIFFLGVLKDL 183

Qy	238	TFIEHGHTLVE--NLINFEKMRMRAAR	265
Db	184	TFIEHGPNPDYLDNTNLINFEKMRMTAKIIR	213

RESULT 3
US-09-922-199A-9
Sequence 8 Application US/00033100A

```

sequence 9, Application 05703922199A
Publication No. US20020187138A1
GENERAL INFORMATION:
APPLICANT: Meyers, Rachel
TITLE OF INVENTION: 15368, A NOVEL HUMAN GTP-RELEASING
TITLE OF INVENTION: FACTOR FAMILY MEMBER AND USES THEREFOR
FILE REFERENCE: 381552002500
CURRENT APPLICATION NUMBER: US/09/922,199A
CURRENT FILING DATE: 2001-08-02
PRIOR APPLICATION NUMBER: 60/222,622
PRIOR FILING DATE: 2000-08-02
NUMBER OF SEQ ID NOS: 28
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 9
LENGTH: 238
TYPE: PR1

```

```

; ORGANISM: Artificial Sequence
;
; FEATURE:
; OTHER INFORMATION: Consensus amino acid sequences
US-09-922-199A-9

```

Query Match	23.8%;	Score 418.5;	DB 10;	Length 238;
-------------	--------	--------------	--------	-------------

Matches	89;	Conservative	38;	Mismatches	56;	Indels	19;	Gaps	4;
---------	-----	--------------	-----	------------	-----	--------	-----	------	----

QY 79 KDLAQQLTDHDMSELSIHQVELIHYV-----LGPQHLDVYTIANIEREMKRFNELQYW 132
::||| || | : || | : || | : ||| ::||| : ||
Db 7 ELALQQLTLDFELFKKIEPSECLGVWMSNREKKGENL----SPNLKEFLORFNNSISYW 62

```

Qy      133 VATELCCLCPVGPRAQLKRFIKLAHLKEQKNLSNFVAWFGLSNALSIRLAHTWERLP 1922
        |||: | ||:: |||: | : | | ||: |||: | | ||::|
Db      63 VATELISSEKPPQRAKTIIEFKVAQHCRELNFNLSMAIYSGNLSSSYRLKKTWEKVP 1222

```

```
QY      193 HKVRKLYSALERLDDPSNNHRYRLAKLS-----PPVPIEMPLLKDQTFIHEGSH 245
        : :: | |::|::| : | : ||::|||
Db      123 KETKLFEELSELMDPSSNYYKNYRELTKSCGFVNHSQPCCIPFLGYLLKDLTFIHEGNP 182
```

```
QY      246 TLVE--NLINFEKMRMARAAR 265
        :: |||||11111:1
Db      183 DYLDNTNLINFEKMRMIKIIR 204
```

RESULT 4
US-09-911-826A-2

```

; GENERAL INFORMATION:
;
; APPLICANT: Rotin, Daniela and Pham, Nam
;
; TITLE OF INVENTION: PAC Activator Nuclo

```

```

; TITLE OF INVENTION: Methods of use
;
; FILE REFERENCE: DDW-5001-US
;
; CURRENT APPLICATION NUMBER: US/09/911,826A
;
; CURRENT FILING DATE: 2002-02-26

```

; PRIOR APPLICATION NUMBER: PCI/CA00/00042
 ; PRIOR FILING DATE: 2000-01-20
 ; PRIOR APPLICATION NUMBER: 2,259,830
 ; PRIOR FILING DATE: 1999-01-20

```

; NUMBER OF SEQ ID NOS. 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 1499

```

ORGANISM: Homo sapiens
US-09-911-826A-2

Query Match	23.3%	Score 409	DB 10	Length 1499
Best Local Similarity	30.9%	Pred. No. 4.2e-33		
Matches 107; Conservative	57	Mismatches 122	Indels 60	Gaps 5

QY 35 VAAISGLMERLFVYNPOEVHEHLRPHRDGPTVGSAAEGDLDVSAKDLGOLTHDMSLEN 94
Db 675 LARIQISGRYYIKNNMETERETLSDDEDQAOLLRQSISLLQSLVEVAALQSLMNFELR 734

QY 95 SIHQVELHYVLGPOHLRVVTT-ANLERMRFRFELQWVATLCLCPVGPRROLIRKE 153
Db 735 NIEPTEYIDFL---KLRSKTSCLNLRREBEVINQETFEVASLELTRFNOLKRKKIKIH 791

QY 154 IKLAHLTEOKNUNTSFAVMEGJNSATSRLAHMERPHKVRRLYSALERLDPSPWNR 213
 792 IKTLAHCECKNNSMFAIISGLALAPARLRTTWKLEPLNKYEKLFODLDDLPSSRNA 851
 Db

QY 214 VVRL--AKLSPPVIPENPLIKDMFTFHEGNTLVLEINFEKRMRAAAMLNHC 271
||| | : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
Db 852 KIRNVLSOMLOPPIIFLFPVKKIDLTFFEGNSKVQLVLEFKRLMAKETIHWGRMA 911

QY 272 SHHPVPLEPLRSR-----VSLLHDS 292
| | |
| | |
Db 912 SVNMDPALMFRTRKKKWSLGSLSQGSTNAVLDVAQTGGHKRRKYVRSSFLNAKKLYEPA 971

OY 293 QVARISTCSQSLSRSPASTWAVYQOLKVIDNORELSRLSRELEP 338
Db 972 QMAR-----KVQYLSNLELEMBEESQTLSTLQCEP 1002

RESULT 5
US-09-822-635-4
; Sequence 4, Application US/09822635
; Patent No. US20010039331A1
; GENERAL INFORMATION:
; APPLICANT: Hunter, John J.
; APPLICANT: Meyers, Rachel A.
; TITLE OF INVENTION: 16836, A NOVEL HUMAN PHOSPHOLIPASE C AND
; TITLE OF INVENTION: USES THEREOF
; FILE REFERENCE: 10448-035001
; CURRENT APPLICATION NUMBER: US/09/822,635
; CURRENT FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: US 60/193,921
; PRIOR FILING DATE: 2000-03-31
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 261
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: consensus sequence
US-09-822-635-4

Query Match 23.1%; Score 405.5; DB 9; Length 261;
Best Local Similarity 37.6%; Pred. No. 7e-34;

Matches 106; Conservative 41; Mismatches 92; Indels 43; Gaps 8;

OY 75 LVSAKDLAQLTLDHMSLFSIHQVELIHVLPQHLD---VTTALERFMRFNLO 131
Db 3 LDPDELEQLTLDLDFELFKIEPSELLGSVWGKRSKSPSPPLAPOLLEAFIERFNVSN 62
OY 132 WVAETELCLCPV--PGPRAQLLRKFIKLAHLKEQKNLSFPAVFGLSNSAISRLAHTWE 189
Db 63 WVAETELKQTTLKPKRAEVLKFIKAKHCRELNNSLMATVSALSSPSISRLKKTWE 122
OY 190 RLPHKVKLYSALERLDP--WNHRYRLAL-----AKLSPYIPMPPLLKDMPTIHE 242
Db 123 KLSKYKRLFEELDELDPSEERFNKRYREALSCNKPVPPLGYLKDLPIDE 182
OY 243 GNTLVEN-----LINFEMKMMARAAMLHCRSH-NPVPLSPRSVSHLEDQVARI 297
Db 183 GNPDELENGTKGLVNFERRKRIKILREIQLOSACOPYMLKPRNDIQELLRAS----- 237
OY 298 STCSQSLSRSPASTWAVYQOLKVI-DNORELSRLSRELEP 338
Db 238 -----RPLEVLPDEEDELVELSLRIEP 259

RESULT 6
US-10-176-306-32
; Sequence 32, Application US/10176306
; Publication No. US20030130465A1
; GENERAL INFORMATION:
; APPLICANT: Meyers, Rachel A.
; APPLICANT: Curtis, Rory A. J.
; APPLICANT: Gluckmann, Maria Alexandra
; APPLICANT: Bandaru, Rajasekhar
; APPLICANT: Kapeller-Libermann, Rosana
; TITLE OF INVENTION: NOVEL HUMAN GENES AND METHODS OF USE THEREOF
; FILE REFERENCE: 10448-195001
; CURRENT APPLICATION NUMBER: US/10/176,306
; CURRENT FILING DATE: 2002-06-20
; PRIOR APPLICATION NUMBER: 10/001,137
; PRIOR FILING DATE: 2001-11-14
; PRIOR APPLICATION NUMBER: PCT/US01/45291
; PRIOR FILING DATE: 2001-11-14
; PRIOR APPLICATION NUMBER: 60/248,362

; PRIOR FILING DATE: 2000-11-14
; PRIOR APPLICATION NUMBER: 60/248,331
; PRIOR FILING DATE: 2000-11-14
; PRIOR APPLICATION NUMBER: 60/248,365
; PRIOR FILING DATE: 2000-11-14
; PRIOR APPLICATION NUMBER: 60/250,077
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: 60/250,327
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: 60/250,176
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: 10/023,617
; PRIOR FILING DATE: 2001-12-18
; PRIOR APPLICATION NUMBER: PCT/US01/49416
; PRIOR FILING DATE: 2001-12-18
; PRIOR APPLICATION NUMBER: 60/256,249
; PRIOR FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: 60/256,405
; PRIOR FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: 10/083,248
; PRIOR FILING DATE: 2001-10-22
; PRIOR APPLICATION NUMBER: PCT/US01/46717
; PRIOR FILING DATE: 2001-10-22
; PRIOR APPLICATION NUMBER: 60/242,324
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/242,518
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/241,989
; PRIOR FILING DATE: 2000-10-20
; NUMBER OF SEQ ID NOS: 86
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 32
; LENGTH: 261
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: consensus sequence
US-10-176-306-32

Query Match 23.1%; Score 405.5; DB 16; Length 261;
Best Local Similarity 37.6%; Pred. No. 7e-34;

Matches 106; Conservative 41; Mismatches 92; Indels 43; Gaps 8;

OY 75 LVSAKDLAQLTLDHMSLFSIHQVELIHVLPQHLD---VTTALERFMRFNLO 131
Db 3 LDPDELEQLTLDLDFELFKIEPSELLGSVWGKRSKSPSPPLAPOLLEAFIERFNVSN 62
OY 132 WVAETELCLCPV--PGPRAQLLRKFIKLAHLKEQKNLSFPAVFGLSNSAISRLAHTWE 189
Db 63 WVAETELKQTTLKPKRAEVLKFIKAKHCRELNNSLMATVSALSSPSISRLKKTWE 122
OY 190 RLPHKVKLYSALERLDP--WNHRYRLAL-----AKLSPYIPMPPLLKDMPTIHE 242
Db 123 KLSKYKRLFEELDELDPSEERFNKRYREALSCNKPVPPLGYLKDLPIDE 182
OY 243 GNTLVEN-----LINFEMKMMARAAMLHCRSH-NPVPLSPRSVSHLEDQVARI 297
Db 183 GNPDELENGTKGLVNFERRKRIKILREIQLOSACOPYMLKPRNDIQELLRAS----- 237
OY 298 STCSQSLSRSPASTWAVYQOLKVI-DNORELSRLSRELEP 338
Db 238 -----RPLEVLPDEEDELVELSLRIEP 259

RESULT 7
US-09-911-826A-7
; Sequence 7, Application US/09911826A
; Patent No. US20020143164A1
; GENERAL INFORMATION:
; APPLICANT: Rotin, Daniela and Pham, Nam
; TITLE OF INVENTION: RAS Activator Nucleic Acid Molecules, Polypeptides and
; FILE REFERENCE: DDW-5001-US

```

: CURRENT APPLICATION NUMBER: US-09/911.826A
: CURRENT FILING DATE: 2002-02-26
: PRIOR APPLICATION NUMBER: PCT/CA00/00042
: PRIOR FILING DATE: 2000-01-20
: PRIOR APPLICATION NUMBER: 2,259,830
: PRIOR FILING DATE: 1999-01-20
: NUMBER OF SEQ ID NOS: 27
: SOFTWARE: PatentIn Ver. 2.1
: SEQ ID NO 7
: LENGTH: 307
: TYPE: prt
: ORGANISM: Drosophila melanogaster
US-09-911-826A-7

```

Query Match	22.5%	Score 396.57	DB 10	Length 307
Best Local Similarity	34.3%	Pred. No. 7,8e-33		
Matches	103	Conservative	54	Mismatches 102; Indels 41; Gaps 9

QY	76	VSADLDAGQLDHDWMSLFNSTHOVELHYUAGPOLRD-VTTANLERMRRENEIQYWA	134
		..:::	
DB	8	LNAVELLIQTLLO--FANPROLESTEVDELLERSRGVPMKSAEYLVNRMEMFVV	64
		..:::	
QY	135	TELICVPGPRAOLDKRFIKLAHLKEOKNLSEFAMFGSNGAISRLATWERLPHK	194
		..: ::: ::: ::: ::: :::	
DB	65	SEICAEHNIVYRMKTIQKOFILAHCKRCRNFMSFEALVSGDHGAASRLQGTWKLKSL	124
		..:::	
QY	195	VKRIYSALERLDDSSMNHRYR-LALAKL--SPYIYPMPLLDKMTFHGNGHTLVNL	251
		..:::	
DB	125	YQRFENLDQDMDSRNMSKYRQVLSAEILAOHPPIYVKKDTFTHGNDYRVGL	184
		..:::	
QY	252	INEFKMKMMAAARMLLHHCGRSHNEVPL-----SPLR-----SRVSHLEDQVARI	297
		::: :::	
DB	185	VNEFKLMLAKRYVRLTHMCS-SPYDLSTIELKGQSPSNLFLSNQMSASQSNNAAGTV	243
		..:::	
QY	298	-----STCEGSLSTRSP-----ASTWAVYQQLKYDNRQELSRUSRELEP	338
		..:::	
	244	IAANAGATTKIKRRKSTAAPMPKMFEEBAQYVRVAKALNSLKLTISDDLLHKFSLEDEP	303

```

RESULT 8
US-09-922-199A-5
: Sequence 5, Application US/09922199A
: Publication No. US20020187138A1
: GENERAL INFORMATION:
: APPLICANT: Meyers, Rachel
: TITLE OF INVENTION: 15368, A NOVEL HUMAN GTP-RELEASING
: TITLE OF INVENTION: FACTOR FAMILY MEMBER AND USES THEREFOR
: FILE REFERENCE: 381552002500
: CURRENT APPLICATION NUMBER: US/09/922,199A
: CURRENT FILING DATE: 2001-08-02
: PRIOR APPLICATION NUMBER: 60/222,622
: PRIOR FILING DATE: 2000-08-02
: NUMBER OF SEQ. ID NOS: 28
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 5
: LENGTH: 227
: TYPE: PRT
: ORGANISM: Artificial Sequence
: FEATURE:
: OTHER INFORMATION: Consensus amino acid sequence
: US-09-922-199A-5

```

Query Match	14.0%;	Score 246.5;	DB 10;	Length 227;
Best Local Similarity	30.28%;	Pred. No. 2.4e-17;		
Matches 68; Conservative	39;	Mismatches 77;	Indels 41;	Gaps 6;

```

QY 75 IYAKDLAGOITBHDMSLFNSHOVELIHYVL-----GPO-----HLDVYTANERMR 124
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2 LDRLPLAKOITLLEHLEFKKIIPTFECIGSMKHEGPOVMCKSLYKCKGEESPINDKIK 61
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 125 REVELQYWA-----TELCLCPGVRPQOLRKFKTLAHLKEOKNLNSSFVFM 173
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 62 NFQOLNFVSEITLKQEPPTNPFRSKVDPKPKRAEVIOKFLQVADHCEMLNPFSLAI 121
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

Oy	17	FGLSNASISLSTALWRLDLHKKVKRL-YSLSLESLDLPSSNNHYRRLATLAKSP-----	224
Db	122	SALYSPSTRLLKKTWKXVPQSLKLIFELNKLMDSDPNFSNYRELLKLSIQLPLKPMLDD	161
Oy	225	-PYIPMPDLLKDMDFIEHGNTLYE-----NLINFEKKR	258
Db	182	DEYRAPCVPEPGYYISDLTFLEEAGNDYLHNKRITNLTAVNSKKR	226

RESULT 9
US-10-176-306-31

```

Sequence 3 Application US/10176306
Publication No. US20030130485A1
GENERAL INFORMATION:
APPLICANT: Meyers, Rachel A.
APPLICANT: Curtis, Rory A. J.
APPLICANT: Glucksmann, Maria Alexandra
APPLICANT: Bandaru, Rajasekhar
APPLICANT: Kapeller-Liberman, Rosana
TITLE OR INVENTION: NOVEL HUMAN GENES AND METHODS OF USE THEREOF
FILE REFERENCE: 10448-195001
CURRENT APPLICATION NUMBER: US/10/176,306
PRIORITY FILING DATE: 2002-06-20
PRIOR APPLICATION NUMBER: 10/001,137
PRIOR FILING DATE: 2001-11-14
PRIOR APPLICATION NUMBER: PCT/US01/45291
PRIOR FILING DATE: 2001-11-14
PRIOR APPLICATION NUMBER: 60/248,362
PRIOR FILING DATE: 2000-11-14
PRIOR APPLICATION NUMBER: 60/248,331
PRIOR FILING DATE: 2000-11-14
PRIOR APPLICATION NUMBER: 60/248,365
PRIOR FILING DATE: 2000-11-14
PRIOR APPLICATION NUMBER: 60/250,077
PRIOR FILING DATE: 2000-11-30
PRIOR APPLICATION NUMBER: 60/250,327
PRIOR FILING DATE: 2000-11-30
PRIOR APPLICATION NUMBER: 60/250,176
PRIOR FILING DATE: 2000-11-30
PRIOR APPLICATION NUMBER: 10/023,617
PRIOR FILING DATE: 2001-12-18
PRIOR APPLICATION NUMBER: PCT/US01/49416
PRIOR FILING DATE: 2001-12-18
PRIOR APPLICATION NUMBER: 60/256,249
PRIOR FILING DATE: 2000-12-18
PRIOR APPLICATION NUMBER: 60/256,405
PRIOR FILING DATE: 2000-12-18
PRIOR APPLICATION NUMBER: 10/083,248
PRIOR FILING DATE: 2001-10-22
PRIOR APPLICATION NUMBER: PCT/US01/6717
PRIOR FILING DATE: 2001-10-22
PRIOR APPLICATION NUMBER: 60/242,324
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/242,518
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/241,989
PRIOR FILING DATE: 2000-10-20
NUMBER OF SEQ ID NOS: 86
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 31
LENGTH: 227
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURES:
OTHER INFORMATION: consensus sequence
US-10-176-306-31
```

Query Match	14.08;	Score 246.5;	DB 16;	Length 227;
Best Local Similarity	30.28;	Pred. No. 2.4e-17;		
Matches 68; Conservative	39;	Mismatches 77;	Indels 41;	Gaps 6

QY 75 LVS AKDLAQ L TDH DWS L FNS I HQV ELIHYL --- GPQ ----- HLRDVT TANLER FMR 124

[illegible]

```

RESULT 10
US-09-940-836A-2
: Sequence 2, Application US/09940836A
: Patent No. US20020146800A1
: GENERAL INFORMATION:
: APPLICANT: Millennium Pharmaceuticals, Inc.
: APPLICANT: Curtiss, Rory
: TITLE OF INVENTION: 48921, A NOVEL HUMAN GTP RELEASING
: TITLE OF INVENTION: FACTOR AND USES THEREFOR
: FILE REFERENCE: 38155-20031.00
: CURRENT APPLICATION NUMBER: US/09/940.836A
: CURRENT FILING DATE: 2001-08-27
: PRIOR APPLICATION NUMBER: US 60/228.760
: PRIOR FILING DATE: 2000-08-30
: NUMBER OF SEQ ID NOS: 12
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 2
: LENGTH: 472
: TYPE: prt
: ORGANISM: Homo sapiens
US-09-940-836A-2

```

Query Match	13.5%	Score 238	DB 10	Length 472
Best Local Similarity	28.5%	Pred. No. 5.6e-16		
Matches	82	Conservative 42	Mismatches 94	Indels 70
			Gaps	10
Qy	1	MAALAOEGRWTKGQVLVKNVNSAGDAIGLQPDARGVATSLGNERLTVV--NPQEVHELIP	58	
		:	:	:
Db	163	LAALSQYE-----EVLAKISSTSS-----IDRLTVLTKPKPSIOR---	196	
Qy	59	HPDOLGPTVGSAGELDLVSANKD---LAGQLTDHDMSLNSIQVELIHVYLPOHLRD-	113	
		:	:	:
Db	197	-----DITVCNDPRTYLAQOLHIELERLNYIGREEFQVAFQKDPDLND	241	
Qy	114	-----VTTANLEFRMRRENELOVATETELCLPVGRPAOILKRFIKLANHLKROKNL	166	
		:	:	:
Db	242	KSCYSERRKTRTLEAVYEWFNRLSYLVAETICMPVKKRRARIEYFIDVARECFNIGNF	301	
Qy	167	NSFPAVMGCLNSAISRLANHTWERLPHKRYKRLXSALERLDDPSMNRYVRRLA-----	219	
		:	:	:
Db	302	NSLMAIISGMNNSPVSRLKKTWAKY--KPAK-FDILEHOMDPSNNYNTALRGAAORS	358	
Qy	220	-----AKLSPPVPIPFMPLLLDKMTTIEHG--NHLTVENLHNEEKRRMA 262		
		:	:	:
Db	359	LTAHSSRRKVIPIPFSLIKDITYLENGGCANRLPNCHINIEFKEFWELAK 406		

```

RESULT 11
US-10-275-173-2
: Sequence 2, Application US/10275173
: Publication No. US20030106932A1
: GENERAL INFORMATION:
: APPLICANT: Merck Patent GmbH
: TITLE OF INVENTION: Ras Gguanidine-nucleotide -exchange factor (NRG1)
: FILE REFERENCE: NRG1CSMs
: CURRENT APPLICATION NUMBER: US/10/275.173
: CURRENT FILING DATE: 2002-11-04

```

```

: NUMBER OF SEQ ID NOS: 4
: SOFTWARE: PatentIn Ver. 2.1
:
: SEQ ID NO 2
:
: LENGTH: 473
:
: TYPE: prt
:
: ORGANISM: Homo sapiens
:
: OS-10-2/5-173-2

```

	Query Match	13.5%	Score 237:	DB 15:	Length 473:
	Best Local Similarity	28.1%:	Pred. No. 7.1e-16:		
	Matches	81:	Conservative 43:	Mismatches 94:	Indels 70: Gaps 10:
OY	1	MAALAEQDGWATGGVLVKNVSAGDAIGLQDPARGVATSLGNLERLFVV--PROEYNELRP	58		
Dd	164	LAAISOYE-----EVLAKISST-----TDRITLVKTQRQSIR----	197		
OY	59	HPDOLGPTVSGAEGLDIVLSAND---LAGOLTIDHWSLEFNSIHVELIHYLGROILRD	113		
Dd	198	-----DIITVCNDPDTLAQOGLTHIELEERLYANTIGREEFQARVQAKRDLND	242		
OY	114	-----VTIANLEFRMRRENELQWATELCISLRVERRAQLLKFKTKLANLKEOKNL	166		
Dd	243	KSCYSERKKTRNLEAVVEFMRLSYATVEICMPYKKKKRRAMRIEEFYDAVARECEFINF	302		
OY	167	NSFEFAVMGDSNASISRLANHWEMLRPHNVKRKLVSALLERLLDRSMNHRYURLAL-----	219		
Dd	303	NSLMAIIISGMMSNPVSRLKTTMAVY--KTA*-FDLEIQMDPPSSMFUYNRALRGCAORS	359		
OY	220	----AKLSRVIPREMRLCLKMTFTIHES-NHTLVEENCLNFPMKRMJAM	262		
Dd	360	LTAASSRKVIYIPFFSLIKDYIFLNBCANRLPRGHVAFKEFWELAK	407		

```

1 RESULT 12
2 US-09-864-761-39774
3 : Sequence 39774, Application US/09864761
4 : Patent No. US20020048763A1
5 :
6 : GENERAL INFORMATION:
7 :   APPLICANT: Penn, Sharron G.
8 :   APPLICANT: Rank, David R.
9 :   APPLICANT: Hanzel, David K.
10 :
11 : APPLICANT: Chen, Wensheng
12 :
13 : TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
14 :   FILE REFERENCE: Aecomica-X-1
15 :   TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
16 :
17 : CURRENT APPLICATION NUMBER: US/09/864,761
18 :
19 : CURRENT FILING DATE: 2001-05-23
20 :
21 : PRIOR APPLICATION NUMBER: US 60/180,312
22 :
23 : PRIOR FILING DATE: 2000-02-04
24 :
25 : PRIOR APPLICATION NUMBER: US 60/207,456
26 :
27 : PRIOR FILING DATE: 2000-05-26
28 :
29 : PRIOR APPLICATION NUMBER: US 09/632,366
30 :
31 : PRIOR FILING DATE: 2000-08-03
32 :
33 : PRIOR APPLICATION NUMBER: GB 24263, 6
34 :
35 : PRIOR FILING DATE: 2000-10-04
36 :
37 : PRIOR APPLICATION NUMBER: US 60/236,359
38 :
39 : PRIOR FILING DATE: 2000-09-27
40 :
41 : PRIOR APPLICATION NUMBER: PCT/US01/00666
42 :
43 : PRIOR FILING DATE: 2001-01-30
44 :
45 : PRIOR APPLICATION NUMBER: PCT/US01/00667
46 :
47 : PRIOR FILING DATE: 2001-01-30
48 :
49 : PRIOR APPLICATION NUMBER: PCT/US01/00664
50 :
51 : PRIOR FILING DATE: 2001-01-30
52 :
53 : PRIOR APPLICATION NUMBER: PCT/US01/00669
54 :
55 : PRIOR FILING DATE: 2001-01-30
56 :
57 : PRIOR APPLICATION NUMBER: PCT/US01/00665
58 :
59 : PRIOR FILING DATE: 2001-01-30
60 :
61 : PRIOR APPLICATION NUMBER: PCT/US01/00668
62 :
63 : PRIOR FILING DATE: 2001-01-30
64 :
65 : PRIOR APPLICATION NUMBER: PCT/US01/00663
66 :
67 : PRIOR FILING DATE: 2001-01-30
68 :
69 : PRIOR APPLICATION NUMBER: PCT/US01/00662
70 :
71 : PRIOR FILING DATE: 2001-01-30
72 :

```

```
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Anomax sequence Listing Engine vers. 1.1
; SEQ ID NO 39774
; LENGTH: 59
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC004241.1
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 7.9
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 6.3
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 6.9
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 7.9
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 7.6
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 8.4
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 7.4
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 7.2
; OTHER INFORMATION: EST_HUMAN HIT: BE379558.1, EVALUE 5.00e-19
; OTHER INFORMATION: SWISSPROT HIT: P48308, EVALUE 6.00e+00
US-09-864-761-39774
```

```
Query Match 13.0%; Score 229; DB 9; Length 59;
Best Local Similarity 100.0%; Pred. No. 2.1e-16;
Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 57 IHPDQIGPTVGSABGLDVLVSADLAGOITDHDMSLFNSIHQV 99
DB 1 IHPDQIGPTVGSABGLDVLVSADLAGOITDHDMSLFNSIHQV 43
```

```
RESULT 13
US-09-805-455-7
; Sequence 7, Application US/09805455
; Publication No. US20030166203A1
; GENERAL INFORMATION:
; APPLICANT: GAN, Weimin
; TITLE OF INVENTION: ISOLATED HUMAN RAS-LIKE PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING THESE HUMAN RAS-LIKE
; FILE REFERENCE: C1001165
; CURRENT APPLICATION NUMBER: US/09/805,455
; CURRENT FILING DATE: 2001-03-14
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 591
; TYPE: PRT
; ORGANISM: Human
US-09-805-455-7
```

```
Query Match 12.6%; Score 221; DB 12; Length 591;
Best Local Similarity 24.8%; Pred. No. 4.7e-14;
Matches 84; Conservative 55; Mismatches 122; Indels 78; Gaps 10;
```

```
QY 26 ISLQDPARGVATSLGNERLFFVNPQEVHELIPHPDQIGPTVGSAGLDLVSADLAGO 85
DB 95 IHWTEEFREVASQLGYEKHVSIDISS--IPSYDMRMRYQRRK---VSKKGACLL 146
QY 86 THDMSLFNSIHQVLIHVLPQH--LDDVT-----TANLERFMRRFE 128
DB 147 FHH-----LEPIELAEHLTFLEKSFRRISFTDYQSYVHGCLENNPTLERSTALFNG 199
QY 129 LQYVATELCLCPVGPQRQLRKFKLAAHLKEQKNLSFPAVMGSLNSAISRLAHW 188
```

```
DB 200 ISKMYQVQNLVSKPTPOQRAEVITKFTINAKKLIQLKNEFTLAAVGGLSHSSISRLKETH 259
QY 189 ERLPHKVRKLYSALERILDPWNHRYRLALAKLSPVLPFMLLKDMFTFHEGHTLV 248
DB 260 SHLSSEVTKNMNMETLVSNGVYKRAFPADCDGCFIPILGVHLKDLIAVH-----VI 314
QY 249 -----ENLINFKKMMMAARAAMLHCHSHNPVLS---PLRSRVSHLHDSQVARIKT 299
DB 315 FPDWTEENKYNVYIKMQLS-----VTLSELVSLQNASHLPEPMDLITLLT 360
QY 300 CSEQSLSTRSPASTWAVYQQLKVIDNQRELSRLRELP 338
DB 361 LS-----LDLYTTEDDIYKLVLVLP 381
```

```
RESULT 14
US-10-177-293-399
; Sequence 399, Application US/10177293
; Publication No. US20030124128A1
; GENERAL INFORMATION:
; APPLICANT: Lillie, James
; APPLICANT: Glatte, Karen
; APPLICANT: Zhao, Xumel
; APPLICANT: Ganavarpu, Manjula
; APPLICANT: Kamalakar, Shubhangi
; APPLICANT: Mertens, Maureen
; APPLICANT: Myer, Vic
; APPLICANT: Wang, Youzhen
; APPLICANT: Xu, Yongyao
; APPLICANT: Hoersch, Sebastian
; APPLICANT: Monahan, John
; APPLICANT: Meyers, Rachel E.
; APPLICANT: Bast Jr., Robert C.
; APPLICANT: Hortobagyi, Gabriel N.
; APPLICANT: Puzsai, Lajos
; APPLICANT: Sahin, Aysegul
; APPLICANT: Mills, Gordon B.
; TITLE OF INVENTION: COMPOSITIONS, KITS, AND METHODS FOR IDENTIFICATION, ASSESSMENT
; TITLE OF INVENTION: PREVENTION, AND THERAPY OF BREAST CANCER
; FILE REFERENCE: MRI-038
; CURRENT APPLICATION NUMBER: US/10/177,293
; CURRENT FILING DATE: 2002-06-21
; PRIOR APPLICATION NUMBER: US 60/299,887
; PRIOR FILING DATE: 2001-06-21
; PRIOR APPLICATION NUMBER: US 60/301,572
; PRIOR FILING DATE: 2001-06-27
; PRIOR APPLICATION NUMBER: US 60/306,501
; PRIOR FILING DATE: 2001-07-18
; PRIOR APPLICATION NUMBER: US 60/325,002
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US 60/362,585
; PRIOR FILING DATE: 2002-03-05
; PRIOR APPLICATION NUMBER: US 60/xxx,xxx
; PRIOR FILING DATE: 2002-05-14
; NUMBER OF SEQ ID NOS: 506
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 399
; LENGTH: 782
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-177-293-399
```

```
Query Match 11.8%; Score 207.5; DB 15; Length 782;
Best Local Similarity 22.7%; Pred. No. 1.9e-12;
Matches 71; Conservative 50; Mismatches 101; Indels 91; Gaps 6;
```

```
QY 51 QEVHELIPHPDQIGPTVGSABGLD-----VSADLAGOITDHDMSLFNSIHQVLI 101
DB 136 EEFQELV-----KAKGEELHCRILDTTQJNARWMSKRLQRIKSSNFKRKRVSL 184
QY 102 IHVYVQPHLRDVT-----ANERFMRRNELEQYVVA 134
```

Db 185 LFQHLPEPEELSEHLTYLEFNSFRRISFSDYQNYLVNSCVKENPTMERSIALCNGISQWQ 244
Qy 135 TELCLCPVGPRAQLLRKFKLAHLKEOKNLNSFPAVMGCLNSAISRLAHTERLPHK 194
Db 245 LMTVSRPTPOLRAEVFTKFTQVAKLQLOLNFNTLMAVIGGLCHSSISRLKETSSHPHE 304
Qy 195 VRKLVSALERLLDPSMNHRYRLAKLSPPVIPPMLLKDMTFIHG----- 243
Db 305 INKVLGEMTELLSSSRNYDNRYRAGECTDEKIPILGVHLKDLISLYEAMPDYLGDKVN 364
Qy 244 -----NHTLVNLINFEKMMARAARMLHHC-----RS 272
Db 365 VHKLLALYNH--ISELVQLQEVAPLEANKDVLHLLTSLDLYYTEDEIYELSYAREPRN 422
Qy 273 HNPVPLSPRSRV 285
Db 423 HRAPPLTPSKPPV 435

RESULT 15

US-09-911-826A-9
: Sequence 9, Application US/09911826A
: Patent No. US20020143164A1
: GENERAL INFORMATION:
: APPLICANT: Rotin, Daniela and Pham, Nam
: TITLE OF INVENTION: RAS Activator Nucleic Acid Molecules, Polypeptides and
: TITLE OF INVENTION: Methods of Use
: FILE REFERENCE: DDW-5001-US
: CURRENT APPLICATION NUMBER: US/09/911,826A
: CURRENT FILING DATE: 2002-02-26
: PRIOR APPLICATION NUMBER: PCT/CA00/00042
: PRIOR FILING DATE: 2000-01-20
: PRIOR APPLICATION NUMBER: 2,259,830
: PRIOR FILING DATE: 1999-01-20
: NUMBER OF SEQ ID NOS: 27
: SOFTWARE: PatentIn Ver. 2.1
: SEQ ID NO 9
: LENGTH: 244
: TYPE: PRT
: ORGANISM: Mus musculus
US-09-911-826A-9

Query Match 11.6%; Score 204; DB 10; Length 244;
Best Local Similarity 25.8%; Pred. No. 7.4e-13;
Matches 60; Conservative 44; Mismatches 115; Indels 14; Gaps 4;

Qy 70 AEGDLVSADKLADQLDHDMSLFNSIHQVELIHYVLGPOHL---RDVTTANLERFMRRF 126
Db 2 AECFETLSAMELAQIILDLHIVFRSIPYEEF---LGQGMKLDKKNERTPYIMKTSQHF 57
Qy 127 NELQYVATELCLCPVGPRAQLLRKFKLAHLKEOKNLNSFPAVMGCLNSAISRLAH 186
Db 58 NEMSNLVASQIMNADISSRPALEKRVAVADICRCLHNYNGVLEITSALNRSPIYRLKK 117
Qy 187 TWERLPHKVRKLYSALERLLDPSMNHRYRLAKLSPPVIPPMLLKDMTFIHGNNHT 246
Db 118 TMAVSVKQTKALMDKLOKTVSSEGRFKMLRETLKNCNPAPVYLGWYLTDLAFLIEEGTPN 177
Qy 247 LV-ENLINFEMKMMARAARMLHHC-----SHNPVPLSPRSRVSHLHDS 292
Db 178 FTEGLVNFSKRMISHIREIROFOQTAVRIDOOPKVIOYLLDKALVIDEDS 230

Search completed: October 7, 2003, 07:25:21
Job time : 67 secs

THIS PAGE BLANK (USPTO)

submitted to the EMBL Data Library, February 1996
A:Reference number: Z19725
A:Accession: T23314
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1305 <WIL>
A:Cross-references: EMBL:Z69664; PIDN:CAA93519.1; GSPDB:GN00022; CESP:T14G10.2
A:Experimental source: clone K04D7
R:WILD, A.
submitted to the EMBL Data Library, January 1996
A:Reference number: Z19954
A:Accession: T24919
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1305 <W12>
A:Cross-references: EMBL:Z68880; PIDN:CAA93100.1; GSPDB:GN00022; CESP:T14G10.2
A:Experimental source: clone T14G10
C:Genetics:
A:Gene: CESP:T14G10.2
A:Map position: 4
A:introns: 450/1; 463/2; 696/2; 763/2; 843/2; 935/3; 1012/1; 1091/1; 1143/1; 1189/2; 129
Query Match
Best Local Similarity 26.1%; Score 459; DB 2; Length 1305;
Matches 116; Conservative 73; Mismatches 130; Indels 38; Gaps 10;
OY 10 WTKQVLYKVNAGDAGTGLDPANGVATSLGLNERLFVYVNPQEVHELIPHPDGLPTVGS 69
DB 662 WSLCECVYITDGVTKQRLPPQEMELERIALNSRYLKNNRSREPLV--PDELAPELLK 719
OY 70 AEGIDIVS--AKDLAGOLTHDMSLFNSHQVELIH--YVGPQHLRVNTANLERFMR 125
DB 720 EAQQLSLNAQVAAQTLTDDFSVSAIEPTEFDNLFKLDKSY---GSPKLEEFQQL 775
OY 126 FNELOYVATVELCLPVPBPRAQLLRKFIKLAHLKEQKNSFFAVMGLNSAISRLA 185
DB 776 FNRKMWVATIEICERHQAOKAKLIKRIKVARCRDLRNNNSMFALMSGLDKAVRRLH 835
OY 186 HTMERLPKRVKLYSALERLLDPSMNRVRYLAKLS--PPVIFPMLLKMTFIHGG 243
DB 836 SSWERVSCKYIRMDIEHQVLDPSPRMSKYRQHLAEQVPPVPIYVVIKKDLTFADG 895
OY 244 NHTLVENINFEKRMARARMLHCHSHNPVPLPSR-----VSHLH--EDSQ 293
DB 896 NATTSSEKLINEKRLKLNKSTIRGVKWL--SSAPYEIASAERSGGVMDALLHMSFENS 954
OY 294 VARI-----STCSQSISTSPASTWAVYVQOLKYIDNQRLEISRLRELEP 338
DB 955 VATRRKMGSKQNPQRKKVYEQALMVRKVS---YLBGLHYVDNEMELDSYDIEP 1008
RESULT 3
JC7736
C3G protein, long type - mouse
C:Species: Mus musculus (house mouse)
C:Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 14-Dec-2001
R:Zhai, B.; Huo, H.; Liao, K.
Biochem. Biophys. Res. Commun. 286, 61-66, 2001
A:Title: C3g, a guanine nucleotide exchange factor bound to adapter molecule c-Crk, has
A:Reference number: JC7736; MUID:21378139; PMID:11485308
A:Contents: 373-11 adpocyte
A:Accession: JC7736
A:Molecule type: mRNA
A:Residues: 1-1086 <ZHA>
A:Cross-references: GB:AF348669
C:Comment: This protein, a guanine nucleotide exchange factor, through the interaction w
ation.
Query Match
Best Local Similarity 16.1%; Score 283.5; DB 2; Length 1086;
Matches 69; Conservative 41; Mismatches 76; Indels 39; Gaps 5;

OY 32 ARGVATSLGLNERLFVYVNPQEVHELIPHPDGLPTVGSABGLDLSAKDLAQLTLDHMS 91
DB 835 ARGVA-----ARPGTLHDFHSH-----EIAEQTLTLDAA 863
OY 92 LENSIHQVELIHYVLPQHLRDVTTANLERFMRFPNELQYVATVELCLPVPBPRAQLLR 151
DB 864 LF---YKIEIEVLLMAKEQNEKSPNLTQTEHFNNKSYVRITIMQEKADREKLL 920
OY 152 KFIKLAHLKEQKNSFFAVMGLNSAISRLAHTWERLPKRVKLYSALERLLDPSMN 211
DB 921 KFIKIMKHLRKLNNFNSYLAISLADSPAIRRL--EMQROTSSEGLAEYCT--LIDSSS 975
OY 212 HAVRYRLAKLSPPVIFPMLLKMTFIHGNNTVENLINEFK 256
DB 976 FRAYRAALSEVEPPCIPYGLILDLDFVHLGNPDYIDGKVNFSK 1020
RESULT 4
T13052
guanine nucleotide exchange factor DC3G - fruit fly (Drosophila melanogaster)
C:Species: Drosophila melanogaster
C:Date: 13-Aug-1999 #sequence_revision 13-Aug-1999 #text_change 17-Nov-2000
C:Accession: T13052
R:Ishimaru, S.; Gaul, U.; Hanafusa, H.
submitted to the EMBL Data Library, March 1998
A:Description: cDNA sequence of Drosophila melanogaster DC3G gene.
A:Reference number: Z17593
A:Accession: T13052
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-1383 <ISH>
A:Cross-references: EMBL:AF053358; NID:g3582764; PID:g3582765; PIDN:AAC35280.1
C:Genetics:
A:Gene: DC3G
A:Cross-references: FlyBase:FBgn0026145
A:Map position: X
C:Superfamily: CDC25-type guanine nucleotide exchange activator homology
F:1147-1376/Domain: CDC25-type guanine nucleotide exchange activator homology <SOS>
Query Match
Best Local Similarity 14.8%; Score 261; DB 2; Length 1383;
Matches 67; Conservative 43; Mismatches 73; Indels 18; Gaps 7;
OY 64 GPYVGSABG-----LDLSAKDLAQLTLDHMSLFNSHQVELIHYVLPQHLRDVTTA 117
DB 1133 GCGSGTAGGQNPISLILDLKSL--ETAEQMTLIDALEFT---KIEIPEVLLAKDCEKSP 1188
OY 118 NLBERFMRFPNELQYVATVELCLPVPBPRAQLLRKFIKLAHLKEQKNSFFAVMGLS 177
DB 1189 NLNKFTEHFNMKSTWANSKIIRLQDAKERKHVKFIKIMKLNKNNYSYLAISALD 1248
OY 178 NSAISRLAHTWERLPKRVKLYSALERLLDPSMNRVRYLAKLSPPVIFPMLLLKD 236
DB 1249 SGPIRRL--EMQKQITEVRSFCA----LIDSSSFAYRQALAEATNPPIPVIYGLILOD 1302
OY 237 MTFIHEGNNH--LVENLINEFK 256
DB 1303 LTFEYVGNODYLSKGVINFSK 1323
RESULT 5
PC1114
SKDC25 protein - yeast (Saccharomyces kluyveri) (fragment)
C:Species: Saccharomyces kluyveri
C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 21-Jul-2000
C:Accession: PC1114
R:Prigoy, T.; Gonzales, E.; Broek, D.
Gene 117, 67-72, 1992
A:Title: Identification and analysis of a DNA fragment from Saccharomyces kluyveri th
A:Reference number: PC1114; MUID:92354938; PMID:1644315
A:Accession: PC1114
A:Molecule type: DNA
A:Residues: 1-1095 <PRI>

A:Cross-references: GB:M82964; NID:q171186; PIDN:AAA34479.1; PID:q171187
 C:Genetics:
 A:Gene: SKDC25
 C:Superfamily: CDC25-type guanine nucleotide exchange activator homology
 C:Keyword: Transmembrane protein
 F:808-1043/Domain: CDC25-type guanine nucleotide exchange activator homology <SOS>

Query Match 12.8%; Score 225; DB 2; Length 1095;
 Best Local Similarity 27.4%; Pred. No. 2, 2e-10;
 Matches 75; Conservative 44; Mismatches 131; Indels 24; Gaps 7;

QY 12 KGVLYK---VNSAGAIQLDPDARGVATSLGNERLFVYNQEVHELIPHPDQLPTVG 68
 DB 752 KGLSMKGNLKNFVPSINSDGSSSTTVY-----PQSSSSVSAVPSSSTTG 800
 QY 69 ----SAGDLVSAKDLAQLTDPHDSLFRNSIHQVELIHVGLPQLRDVTTANLERFAR 124
 DB 801 FRRRLKLLD-IDSLYAKQLTIKEHSLFYKISPFCDLRTWGNKCNMGSKNTEFIS 859
 QY 125 RFEQLQVWATELCLCPVPGPRAQLRKFKLAHLKEQKLNLSFFAVMFGLSAISRL 184
 DB 860 NSHNLNIVYSFMYKQTDIKKRIQLQFNLVAAHCHLNNSLTAIIISALYSPIYRL 919
 QY 185 AHTWERLPKRVKRLYSALERLDPNHNRYRLALAKLSP-PYIPMPMLLKDMTEIHG 243
 DB 920 KRTMAVVPBEYKLLBELNLTMSAKNFIRYQLKLSIGDFCVPFQVGLSDLTFTANG 979
 QY 244 NHTLVN---LINF-EKMRMAAAMHHCRRH 273
 DB 980 NPDLHRNTVLVNFGRKVRILEIKELISYQRSH 1013

RESULT 6

S28098
 guanine-nucleotide releasing factor, Ste6p - fission yeast (Schizosaccharomyces pombe)

C:Species: Schizosaccharomyces pombe
 C:Date: 17-Apr-1993 #sequence_revision 17-Apr-1993 #text_change 01-Dec-2000
 C:Accession: S28098; T40999; T40966
 R:Hughes, D.A.; Fukui, Y.; Yamamoto, M.
 Mature 344, 355-357, 1990
 A:Title: Homologous activators of ras in fission and budding yeast.
 A:Reference number: S28098; MOID:90190870; PMID:2107403
 A:Accession: S28098
 A:Molecule type: DNA
 A:Residues: 1-911 <HUG>
 A:Cross-references: EMBL:X53254; NID:95100; PIDN:CAA37345.1; PID:95101
 R:Lyne, M.; Rajandream, M.A.; Barrell, B.G.; Volckaert, G.
 submitted to the EMBL Data Library, March 1999
 A:Reference number: 221962
 A:Accession: T40999
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-911 <LYN>
 A:Cross-references: EMBL:AL049559; PIDN:CA940184.1; GSPDB:GN00068; SPDB:SPCC1450.17
 R:Lyne, M.; Rajandream, M.A.; Barrell, B.G.; Volckaert, G.
 submitted to the EMBL Data Library, October 1998
 A:Reference number: 221961
 A:Accession: T40966
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 650-911 <LY2>
 A:Cross-references: EMBL:AL031966; PIDN:CAA21435.1; GSPDB:GN00068; SPDB:SPCC1442.01
 A:Experimental source: strain 972n-; cosmid c1442
 C:Genetics:
 A:Gene: ste6; SPCC1450.17; SPDB:SPCC1442.01
 A:Map position: 3
 C:Superfamily: CDC25-type guanine nucleotide exchange activator homology
 F:659-894/Domain: CDC25-type guanine nucleotide exchange activator homology <SOS>

Query Match 11.9%; Score 208.5; DB 2; Length 911;
 Best Local Similarity 26.2%; Pred. No. 4e-09;
 Matches 63; Conservative 43; Mismatches 107; Indels 27; Gaps 5;

QY 51 QEVHELIPH--PDQLPTVGSAGDLVSAKDLAQLTDPHDSLFRNSIHQVELIHVGLP 108
 DB 638 QLEDELSTLNSPD---PIYKDELVLRLPRELAKDLCLERQSFHISRIQELTMD- 693
 QY 109 QHLRDVTTANLERFMR-----FNEQLQVWATELCLCPVPGPRAQLRKFKLAHL 160
 DB 694 -----NLNFSPEKSTFYLNSHLNVFETETIVQEEPRRTNVLAVPIQCDYL 744
 QY 161 KEKKNLSFFAVMFGLSAISRLAHTWERLPKRVKRLYSALERLDPNHNRYRLALA 220
 DB 745 RELNPNASLFSIISALNSSPIHRLRTWNLNLSKTLASPELLNLTLEARKNESYDCE 804
 QY 221 KLSPPYIPMPMLLKDMTEIHGCHNHLVENTLFEKMRMAARAARLHHCRRS----HNPV 276
 DB 805 NCVLPCVPLGLYFTDLTLTKGNKNFQNMIMFKRTKVTYRLINEIKKFGQSGVYMPNP 864

RESULT 7

T21321
 hypothetical protein F25B3.3 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans
 C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 02-Sep-2000
 C:Accession: T21321
 R:Gardner, A.
 submitted to the EMBL Data Library, April 1996
 A:Reference number: Z19406
 A:Accession: T21321
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-638 <ML>
 A:Cross-references: EMBL:Z70752; PIDN:CAA94755.1; GSPDB:GN00023; CESP:F25B3.3
 C:Genetics:
 A:Gene: CESP:F25B3.3
 A:Map position: 5
 A:Interons: 28/1; 66/3; 125/2; 143/3; 171/3; 202/3; 229/3; 333/2; 419/3; 601/2
 C:Superfamily: protein kinase C zinc-binding repeat homology
 F:521-570/Domain: protein kinase C zinc-binding repeat homology <KN>

Query Match 11.7%; Score 206.5; DB 2; Length 638;
 Best Local Similarity 25.9%; Pred. No. 3, 6e-09;
 Matches 81; Conservative 48; Mismatches 129; Indels 55; Gaps 9;

QY 47 VVNPQEVHELIPHPDQLPTVGSAGDLVSAK-----DLAQLTDPHDM 90
 DB 136 VRRPLAKQITVAVDFETLPPTGTPRRPLASKKFLTASLSFVQASPSDISTSLSHIDV 195
 QY 91 SLFNSIHQVELIHVGLPQHLRDVTTANLERFMRFRNEQLQVWATELCLCPVGPRAQL 150
 DB 196 RVLSTRISITELKQYV-KDGLR--SCPMLERSISVFNLSNWOCMLNKTTPKEAREIL 252
 QY 151 RKFKLAHLKEQKLNLSFFAVMFGLSAISRLAHTWERLPKRVKRLYSALERLDPSPW 210
 DB 253 VKFVHAHKLRLKINNFTLMSVGGITHSSVAVRLAKTYAVLSNDIKKELTQTLNLSAQH 312
 QY 211 NHRVYRLALAKLSPV-IPMPMLLKDMTEIHGCHNHLVENTLFEKMRMAARAARLH 269
 DB 313 NFCEYRAAGACKKRRIPITIGVHLKDLVAINCSG-----ANFETKQIS----- 357
 QY 270 CRSHNPVPLSPRSRVSHLEDISOVARISTCSQSLSTRSPASTWYVQOLKY-ID--N 325
 DB 358 --SDKLVLKLSKLLSNFLVFNQKH-----NLPEKNMDLINTLKYSLDIRYN 401
 QY 326 QRELRLSRLEP 338
 DB 402 DDDIYELSLRREP 414

RESULT 8

S29083
 guanine-nucleotide-releasing protein - rat
 N:Alternate names: CDC25 type protein homology

C:Accession: A38985; A46199; I58371
R:Wei, W.; Brock, D.
submitted to GenBank, December 1994
A:Description: Cloning and analysis of the full length human cdc25 CDNA, a ras-specific
A:Accession: A38985
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-1275 <WE1>
A:Cross-references: GB:I26584; NID:94337379; PIDN:AAA58417; PID:9433720
R:Wei, W.; Mosteller, R.D.; Saayal, P.; Gonzales, E.; McKinney, D.; Dasgupta, C.; Li, P.
Proc. Natl. Acad. Sci. U.S.A. 89, 7100-7104, 1992
A>Title: Identification of a mammalian gene structurally and functionally related to the
A:Reference number: A46199; MUID:92357779; PMID:1379731
A:Accession: A46199
A:Molecule type: mRNA
A:Residues: 1047-1054, 'A', 1056-1112, 'G', 1114-1135, 'C', 1137-1275 <WE2>
A:Experimental source: brain
A:Note: sequence extracted from NCBI backbone (NCBIN:111098, NCBIPI:111099)
R:Schweighofer, F.; Faure, M.; Fath, I.; Chevallier-Multon, M.C.; Aplou, F.; Dutrillaux
Oncogene 8, 1477-1485, 1993
A>Title: Identification of a human guanine nucleotide-releasing factor (H-GRF5) specific
A:Reference number: I58371; MUID:93275641; PMID:7684828
A:Accession: I58371
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 787-1275 <RES>
A:Cross-references: GB:I582035; NID:9386046; PIDN:AMB26881.1; PID:9386047
C:Superfamily: CDC25-type guanine nucleotide exchange activator homology; CDC4 homology
P:240-446/Domain: CDC24 homology <CD24>
P:1036-1272/Domain: CDC23-type guanine nucleotide exchange activator homology <SCS>

```

Query Match          11.3% Score 199: DB 2: Length 1275:
Best Local Similarity 27.8%; Pred. No. 4e-08;
Matches      55: Conservative    41: Mismatches    98: Indels        4: Gaps       3:

Oy      70 AEGDLVLSAKMDIAGQLTDHDSLSPNSIHQVELIHVLGPQHL-RDYTTANLERPMRRFNE 128
            || : : : | | | : | : | : | : | : | : | : | : | : | : | :
Db      1033 AEPPENHSALIEAIDLTLDDHLVFVKKIPYEEF--FGGGMKMLEKNERTPYIMKTTHKHND 1090
Oy      129 LQYWATPILCLCPYPGPAQLLRKFIKLAHLKEQQKNUNSFVAFMFLGSNALSIRLAHTW 168
            : : : : : | : : : | : : : | : : : | : : : | : : : | : : : |
Db      1091 ISNLIASIELINEDIIMNRVASIERKWAVADICRCLHNNAVALETTSSNNRSALIFRLKKTW 1150
Oy      189 ERLPHKVYKLTSALEERLDPSMNRHYRLALAKLSPPIPEMPILLKDMRTFHEGGNTLY 248
            : : : : : | : : : | : | : | : : : | : : : | : : : | : : : |
Db      1151 LKYSKQTALLDKIQKLVSSEGRFNKRLREALKNDPCPVILGMTLDIAETEGTENTY 1210
Oy      249 EN-LINEFKMRMMARAAR 265
            | : | : | | | | : |
Db      1211 EDGLVNFSGKMRIISHIR 1228

RESULT 12
S30356
CDC25 protein homolog - yeast (Candida albicans)
C.Species: Candida albicans
C.Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 29-Oct-1999
C.Accession: S30356
R.Goldberg, D.; Matbach, I.; Gross, E.; Levitzki, A.; Simchen, G.
Eur. J. Biochem. 213, 195-204, 1993
A.Title: A Candida albicans homolog of CDC25 is functional in Saccharomyces cerevisiae.
A.Reference number: S30356; MUID:93238685; PMID:8477693
A.Accession: S30356
A.Status: Preliminary
A.Molecule type: DNA
A.Residues: 1-1333 <EGL>
A.Cross-references: KGBL,M94160; NID:g170834; PIDN:AAA34329.1; PID:g170835
C.Superfamily: SH3 homology; CDC25-type guanine nucleotide exchange activator homology
F.37-89/Domain: SH3 homology <SH3B>
F.1064-1305/Domain: CDC25-type guanine nucleotide exchange activator homology <SOS>

Query Match          11.2% Score 196.5: DB 2: Length 1333:

```

```

Best Local Similarity 27.9%, Pred. No. 6.9e+08;
Matches 63; Conservative 45; Mismatches 99; Indels 19; Gaps 6;

QY      -43  ERLRYVNVQGEVHNLPHNDQLGP-TVGA-----EQLDLSAKDLAQLGTLTHDMSLFN 94
      |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db      1033  ERLVY---EKDPYIPNTPKPPAPRLKGLSLKKRPVMDIVY---ELARQLTLLEFKLYC 1085S
      |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY      95  SIHVELIHYVLGFOHLENDVTANLEREMRRRNELQYVWATELCPLVPGPRAOLLRKFI 154
      |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db      1086  KITPFACIACAKWVGKSKSGLSIESIDSTIGCIKANSQNLTFNPGVMIIRKADPPKRVQIIIRFI 1145S
      |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY      155  KLAHLIKKQKMINSPFAMVGLSNASIRLATHWELRPHKRYKLYSALERLIDPSMNIWRY 214
      |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db      1146  QVADKCRQYNNFSSMTATISALYSSPIRLAKTKWTEYVMAADLNLKNNKNNKLNSSRNENE 1205S
      |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY      215  YRLALAKL-SPPVIRFEMPLIKMDGTIEGHNTLVEN---LIIPEK 256
      |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db      1206  YRDVLKFTGSEPCVFCVGLYSDLTFTVHGPRDYLIANTTROYNRAK 1251
      |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

```

```

RESULT 13
Ras guanine nucleotide exchange factor son-of-sevenless (sos) 1 - mouse
S25716
N:Alternate names: probable ras activator
C:Species: Mus musculus (house mouse)
C:Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 05-Nov-1999
C:Accession: S25716; S21991
R:Bowtell, D., Fu, P., Simon, M., Senior, P.
Proc. Natl. Acad. Sci. U.S.A. 89, 6511-6515, 1992
A>Title: Identification of murine homologues of the Drosophila Son of sevenless gene:
F:455-561/Dominant pleckstrin repeat homology <PLK>
A:Reference number: S25714; MUID:92353528; PMID:1621150
A:Accession: S25716
A:Molecule type: mRNA
A:Residues: 1-1336 <DOM>
A:Cross-references: EMBL:Z11574; NID:954134; PIDN:CAA7662.1; PID:954135
C:Superfamily: CDC25-type guanine nucleotide exchange activator homology; plectstrln
F:455-561/Dominant pleckstrin repeat homology <PLK>
F:793-1036/Dominant CDC25-type guanine nucleotide exchange activator homology <SO>

Query Match      11.0%; Score 193; DB 2; Length 1336;
Best Local Similarity 27.9%; Pred. No. 1.3e+07;
Matches    60; Conservative   46; Mismatches    83; Indels     26; Gaps     9;

QY      65  PTV-----GSAEGLDLVSAR--DLGGQLTDHDMSLFNSIHQVELHYVLAQPQHLDVT 115
          |||           | :||:         ::|||       |:|::  |::  |::  ::
DB      776  PTVEHHSIRPCHIEFPDLLTHPELRQLTLESDDLRAVQPSELVGSWTKED-KEIN 834
          :|::  :|::  :|::  :|::  :|::  :|::  :|::  :|::  :|::  :|::  :|::
QY      116 TANLERERKRNELOQVVATVETLCCLPYRG--PRAQLRKPKIKAAHLKEOKNLNSEPAVM 173
          :|::  :|::  :|::  :|::  :|::  :|::  :|::  :|::  :|::  :|::  :|::
DB      835 SPNLKMIKHRTNTTLTWFE-EKCIVETENLEERYAVVSRITIELQVPQELINFNGVLLEV 892
          :|::  :|::  :|::  :|::  :|::  :|::  :|::  :|::  :|::  :|::  :|::
QY      174 FGLSNSAISRLAHMWERLPKHKVRKYLSALERTLDSPSNHHRYRALAKL---SPPVIPFM 230
          :::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::
DB      893 SAMNSPPVRIHDRHFEOIPIRSROKKI---LEEANHLISDHI-YKKYLAKLINSINPCVPFF 947
          :|::  :|::  :|::  :|::  :|::  :|::  :|::  :|::  :|::  :|::  :|::
QY      231 PLLLKDMTFIEHGHTLV---ENLIINFEKRMMA 261
          :|::  :|::  :|::  :|::  :|::  :|::  :|::  :|::  :|::  :|::  :|::
DB      948 GIYLTNILKTGEGNPEVLYRRHGKELINFSKRRRAY 982
          :|::  :|::  :|::  :|::  :|::  :|::  :|::  :|::  :|::  :|::  :|::

RESULT 14
A37488
Ras guanine nucleotide exchange factor son-of-sevenless (sos) 1 - human
C:Species: Homo sapiens (man)
C:Date: 03-Mar-1994 #sequence_revision 18-Nov-1994 #text_change 18-Mar-1997
C:Accession: A37488
R:Chardin, P.; Camonats, J.H.; Gale, N.W.; van Aeist, L.; Schlessinger, J.; Wigler, M.
Science 260, 1338-1343, 1993
A>Title: Human Sos1: a guanine nucleotide exchange factor for Ras that binds to GRB2.
A:Reference number: A37488; MUID:99262494; PMID:8493579
A:Accession: A37488
A>Status: Preliminary; not compared with canonical translation
A:Molecule type: mRNA

```


GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 7, 2003, 07:13:47 ; Search time 25 seconds
(without alignments)
393.787 Million cell updates/sec

Title: US-09-856-679-2

Perfect score: 1759
Sequence: 1 MALALEDGWTGKGLVKNV.....OLKVIDNORELSRLSELEP 338

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 136899 seqs, 29126274 residues

Total number of hits satisfying chosen parameters: 136899

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending-Patents_AA_New:*
1: /cgn2_6/ptodata/1/pae/PCT_NEW_COMB.pep.*
2: /cgn2_6/ptodata/1/pae/US06_NEW_COMB.pep.*
3: /cgn2_6/ptodata/1/pae/US07_NEW_COMB.pep.*
4: /cgn2_6/ptodata/1/pae/US08_NEW_COMB.pep.*
5: /cgn2_6/ptodata/1/pae/US09_NEW_COMB.pep.*
6: /cgn2_6/ptodata/1/pae/US10_NEW_COMB.pep.*
7: /cgn2_6/ptodata/1/pae/US60_NEW_COMB.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed.
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1392.5	79.2	819	1 PCT-US03-28227-3753	Sequence 3753, Ap
2	863.5	49.1	980	1 PCT-US03-28227-3136	Sequence 3136, Ap
3	637.5	36.2	456	6 US-10-435-696-97	Sequence 97, Appl
4	409	23.3	1499	7 US-60-480-890-348	Sequence 348, Appl
5	205.5	11.7	557	7 US-60-483-917-76	Sequence 76, Appl
6	193	11.0	1356	7 US-60-487-610-1523	Sequence 1523, Ap
7	193	11.0	1398	7 US-60-487-610-1524	Sequence 1524, Ap
8	188.5	10.7	1332	7 US-60-487-610-2054	Sequence 2054, Ap
9	182	10.3	699	6 US-10-240-145A-86	Sequence 86, Appl
10	182	10.3	715	6 US-10-240-145A-85	Sequence 85, Appl
11	147.5	8.4	821	6 US-10-296-115-1212	Sequence 112, Ap
12	108	6.1	814	7 US-60-487-610-1754	Sequence 1754, Ap
13	101.5	5.8	437	6 US-10-425-114A-46385	Sequence 46385, A
14	99	5.6	473	1 PCT-US03-22819-2	Sequence 2, Appl
15	94.5	5.4	438	6 US-10-425-114A-40595	Sequence 40595, A
16	88	5.0	569	6 US-10-425-114A-64070	Sequence 64070, A
17	87.5	5.0	2816	6 US-10-240-145A-145	Sequence 145, App
18	87	4.9	271	5 US-09-897-516A-4785	Sequence 4785, Ap
19	87	4.9	2432	1 PCT-US03-28227-4002	Sequence 4002, Ap
20	87	4.9	2456	1 PCT-US03-28227-3997	Sequence 3997, Ap
21	87	4.9	2466	1 PCT-US03-28227-3996	Sequence 3996, Ap
22	87	4.9	2487	1 PCT-US03-28227-4001	Sequence 4001, Ap
23	87	4.9	2497	1 PCT-US03-28227-4000	Sequence 4000, Ap
24	87	4.9	2508	1 PCT-US03-28227-3999	Sequence 3999, Ap
25	87	4.9	2521	1 PCT-US03-28227-3998	Sequence 3998, Ap
26	85	4.8	533	6 US-10-425-114A-51223	Sequence 51223, A

27	84.5	4.8	382	5 US-09-920-923B-5	Sequence 5, Appl1
28	84.5	4.8	408	7 US-60-450-890-2256	Sequence 2256, Ap
29	84.5	4.8	760	7 US-60-490-890-2636	Sequence 2636, Ap
30	84	4.8	386	6 US-10-425-114A-59796	Sequence 59796, A
31	84	4.8	484	1 PCT-US03-10308A-29	Sequence 29, Appl
32	83.5	4.7	333	6 US-10-425-114A-51680	Sequence 51680, A
33	83	4.7	403	5 US-09-897-516A-7280	Sequence 7280, Ap
34	83	4.7	546	5 US-09-897-516A-4950	Sequence 4950, Ap
35	83	4.7	858	7 US-60-482-992-13	Sequence 13, Appl
36	83	4.7	926	7 US-60-480-890-81	Sequence 81, Appl
37	83	4.7	926	7 US-60-489-772-14	Sequence 14, Appl
38	83	4.7	926	7 US-60-498-106-2	Sequence 2, Appl1
39	83	4.7	1376	7 US-60-479-073-490	Sequence 490, App
40	82	4.7	481	7 US-60-487-610-2123	Sequence 2123, Ap
41	82	4.7	576	7 US-60-487-610-2124	Sequence 2124, Ap
42	81.5	4.6	311	6 US-10-425-114A-59795	Sequence 59795, A
43	81.5	4.6	530	6 US-10-425-114A-40263	Sequence 40263, A
44	81	4.6	603	6 US-10-415-656-32	Sequence 32, Appl
45	80.5	4.6	477	6 US-10-425-114A-65276	Sequence 65276, A

ALIGNMENTS

RESULT 1
PCT-US03-28227-3753
Sequence 3753, Application PC/TUS0328227
GENERAL INFORMATION:
: APPLICANT: INCYTE CORPORATION; SCHMIDT, Jeanette P.;
: APPLICANT: WRIGHT, Rachel J.; BRUNS, Christopher M.;
: APPLICANT: MARJANOVIC, Mirjana M.; SHEN, Fan;
: APPLICANT: HARTSHORNE, Tolnette A.; SUCHOROLSKI, Martin;
: APPLICANT: ALTUS, Christina M.; PITTS, Steven J.;
: APPLICANT: ELDER, Linda V.; MOONEY, Elizabeth M.;
: APPLICANT: DELEGANE, Angelo M.; PANESAR, Iqbal S.;
: APPLICANT: BANYVILLE, Steven C.; REDDY, Thirupathi P.;
: APPLICANT: STEVENS, Kristian A.; BLANCHARD, John L.;
: APPLICANT: PANZER, Scott R.; WANG, Xinbao;
: APPLICANT: AU, Alan P.; GERSTIN, Edward H., Jr.;
: APPLICANT: PERALTA, Careyna H.; ANDERSON, Scott E.;
: APPLICANT: RIOUX, Pierre; SHEN, Edward J.;
: APPLICANT: WU, Mingham C.; STUVE, Laura L.;
: APPLICANT: LAGACE, Robert E.; SPIRO, Peter A.;
: APPLICANT: STEWART, Elizabeth A.; WINGROVE, James A.;
: APPLICANT: VIRT, Ursula A.; KIRTON, Edward;
: APPLICANT: XU, Yuming; KWONG, Mary;
: APPLICANT: POLICKY, Jennifer L.; HURWITZ, Bonnie L.;
: APPLICANT: MA, Yan; JACKSON, Jennifer L.;
: APPLICANT: GIETZEN, Darryl; PATURY, Srikanth;
: APPLICANT: SHI, Xiaobing; SUAREZ, Charlyn J.
: TITLE OF INVENTION: MOLECULES FOR DIAGNOSTICS AND THERAPEUTICS
: FILE REFERENCE: PN-0100 PCT
: CURRENT APPLICATION NUMBER: PCT/US03/28227
: CURRENT FILING DATE: 2003-09-12
: PRIOR APPLICATION NUMBER: US 60/410,260
: PRIOR FILING DATE: 2002-09-12
: PRIOR APPLICATION NUMBER: US 60/410,259
: PRIOR FILING DATE: 2002-09-12
: NUMBER OF SEQ ID NOS: 5444
: SOFTWARE: PERL Program
: SEQ ID NO 3753
: LENGTH: 819
: TYPE: PRT
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: misc-feature
: OTHER INFORMATION: Incyte ID No: 932640.PT97P
PCT-US03-28227-3753

Query Match 79.2% Score 1392.5; DB 1; Length 819;
Best Local Similarity 82.0% Pred. No. 1.6e-92;
Matches 277; Conservative 0; Mismatches 0; Indels 61; Gaps 1;

```
QY 1 MALLAOEDGWTGQVLYVKNVNSAGDAIGLOPDARGVATSLGLENRLFVVPQVEHLEIHP 60
DB 543 MALLAOEDGWTGQVLYVKNVNSAGDAIGLOPDARGVATSLGLENRLFVVPQVEHLEIHP 602
QY 61 DOLGPTVGSAGEGLDVSADLADGOLTDHWSLFNSIHQVELLHYVLGPOHLADVTANLE 120
DB 603 DOLGPTVGSAGEGLDVSADLADGOLTDHWSLFNSIHQVEHLEIHP 640
QY 121 RFRMRRENLQVWVATELCLCPVGPRAQLLRKFILAAHLKBOKNLNSFFAVMGISNSA 180
DB 641 -----LKEOKNLNSFFAVMGISNSA 661
QY 181 ISRLAHTWERLPHKVRKLYSALERLDPSMNRVRYRLAKLSPPVYIPMPMLLKDMFTI 240
DB 662 ISRLAHTWERLPHKVRKLYSALERLDPSMNRVRYRLAKLSPPVYIPMPMLLKDMFTI 721
QY 241 HEGNHTLVENLINFERRMARAARLHCRSHNPVPLSRVSHLHEDSOVARISNC 300
DB 722 HEGNHTLVENLINFERRMARAARLHCRSHNPVPLSRVSHLHEDSOVARISNC 781
QY 301 SEOSLSTRSPASTWAVVQOLKVIDNORELSTRLELEP 338
DB 782 SEOSLSTRSPASTWAVVQOLKVIDNORELSTRLELEP 819
```

RESULT 2
PCT-US03-28227-3136

; Sequence 3136, Application PC/TUS0328227

; GENERAL INFORMATION:

; APPLICANT: INCYTE CORPORATION; SCHMIDT, Jeanette P.;
; APPLICANT: WRIGHT, Rachel J.; BRUNS, Christopher M.;
; APPLICANT: MARANOVIĆ, Mirjana M.; SHEN, Fan;
; APPLICANT: HARTSHORNE, Toinette A.; SUCHOROLSKI, Martin;
; APPLICANT: ALTUS, Christina M.; PITTS, Steven J.;
; APPLICANT: ELDER, Linda V.; MOONEY, Elizabeth M.;
; APPLICANT: DELEGEANE, Angelo M.; PANESAR, Iqbal S.;
; APPLICANT: BANYILLE, Steven C.; REDDY, Thirupathi P.;
; APPLICANT: STEVENS, Kristian A.; BLANCHARD, John L.;
; APPLICANT: PANZER, Scott R.; WANG, Xinhao;
; APPLICANT: AU, Alan P.; GERSTIN, Edward H., Jr.;
; APPLICANT: PERALTA, Careyna H.; ANDERSON, Scott E.;
; APPLICANT: RIOUX, Pierre; SHEN, Edward J.;
; APPLICANT: WU, Mingham C.; STUVE, Laura L.;
; APPLICANT: LAGACE, Robert E.; SPIRO, Peter A.;
; APPLICANT: STEWART, Elizabeth A.; WINGROVE, James A.;
; APPLICANT: VITT, Ursula A.; KIRTON, Edward;
; APPLICANT: XU, Yuming; KWONG, Mary;
; APPLICANT: POLICKY, Jennifer L.; HUBWITZ, Bonnie L.;
; APPLICANT: MA, Yan; JACKSON, Jennifer L.;
; APPLICANT: GIETZEN, Darryl; PATURY, Srikanth;
; APPLICANT: SHI, Xiaobing; SUIREZ, Charlyn J.
; TITLE OF INVENTION: MOLECULES FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: PN-0100 PCT
; CURRENT APPLICATION NUMBER: PCT/US03/28227
; CURRENT FILING DATE: 2003-09-12
; PRIOR APPLICATION NUMBER: US 60/410,260
; PRIOR FILING DATE: 2002-09-12
; PRIOR APPLICATION NUMBER: US 60/410,259
; PRIOR FILING DATE: 2002-09-12
; NUMBER OF SEQ ID NOS: 5444
; SOFTWARE: PERL Program
; SEQ ID NO 3136
; LENGTH: 980
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No: 928293.PT40P
; NAME/KEY: unsure
; LOCATION: (1) ... (980)
; OTHER INFORMATION: unknown or other
PCT-US03-28227-3136

Query Match 49.1%; Score 863.5; DB 1; Length 980;
Best Local Similarity 50.0%; Pred. No. 1,3e-54;
Matches 169; Conservative 63; Mismatches 81; Indels 25; Gaps 3;

```
QY 1 MALLAOEDGWTGQVLYVKNVNSAGDAIGLOPDARGVATSLGLENRLFVVPQVEHLEIHP 60
DB 665 ISAVADKLGGGLIIVKSSGGEKVYLPNDVSVFTTLTIGRLFACPREDFSLTLP 724
QY 61 DOLGPTVGSAGEGLDVSADLADGOLTDHWSLFNSIHQVELLHYVLGPOHLADVTANLE 120
DB 725 EOBGPTVGVTFPELMSSDGLAYQMTIYMEFLPCVHELELLYHFG-RHNNKTTAND 783
QY 121 RFRMRRENLQVWVATELCLCPVGPRAQLLRKFILAAHLKBOKNLNSFFAVMGISNSA 180
DB 784 LFRRENLQVWVATELCLCSQSLSKRVOLLRKFILAAHCKEYKNLNSFFAVMGISNSA 843
QY 181 ISRLAHTWERLPHKVRKLYSALERLDPSMNRVRYRLAKLSPPVYIPMPMLLKDMFTI 240
DB 844 VSRLLATWBEKLSKKEFAEFESLMDPSRNRAVRYRLAKLEPLIPMPMLLKDMFTI 903
QY 241 HEGNHTLVENLINFERRMARAARLHCRSHNPVPLSRVSHLHEDSOVARISNC 300
DB 904 HEGNHTLVENLINFERRMARAARLHCRSHNPVPLSRVSHLHEDSOVARISNC 953
QY 301 SEOSLSTRSPASTWAVVQOLKVIDNORELSTRLELEP 338
DB 954 -----SVRQNLVINDNOFTLSQSHRLEP 977
```

RESULT 3
US-10-435-696-97

; Sequence 97, Application US/10435696

; GENERAL INFORMATION:

; APPLICANT: Wirtz, Ralph
; APPLICANT: Munnes, Marc
; APPLICANT: Kallabis, Harald
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE PREDICTION, DIAGNOSIS, PROGNOSIS
; FILE REFERENCE: Lea 36 108
; CURRENT APPLICATION NUMBER: US/10/435,696
; CURRENT FILING DATE: 2003-05-09
; PRIOR APPLICATION NUMBER: EP03003112.4
; PRIOR FILING DATE: 2003-02-13
; PRIOR APPLICATION NUMBER: EP02010291.9
; PRIOR FILING DATE: 2002-05-21
; NUMBER OF SEQ ID NOS: 314
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 97
; LENGTH: 456
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-435-696-97

Query Match 36.2%; Score 637.5; DB 6; Length 456;
Best Local Similarity 42.0%; Pred. No. 9,1e-39;
Matches 136; Conservative 60; Mismatches 99; Indels 29; Gaps 6;

```
QY 15 VLVKNSAGDAIGLOPDARGVATSLGLENRLFVVPQVEHLEIHPDOLGPTVGSAGEGLD 74
DB 157 ILVAVSSGSEKVLDPTEDCVFTALGINSHPACTRDSYEALVPLPEELVSGPOTE-IH 215
QY 75 LVSAKDLAGOLTDHWSLFNSIHQVELLHYVL-GPOHLADVTANLEFRMRRENLQVW 133
DB 216 RVEPBDVANHLTAFLMEFLRCVHELEFVDYVHGRGRR--TANLELLQGCSEVTHV 273
QY 134 ATELCVPGPRGRLKRFILAAHLKBOKNLNSFFAVMGISNSAISRLAHTWERLPH 193
DB 274 ATEVLCEAPGRADLLKFKIKIAICQONDLFSFVAVMGIDNAVASRLRLTWKELPG 333
QY 194 KVRKLYSALERLDPSMNRVRYRLAKLSPPVYIPMPMLLKDMFTIHEGNHTLVENLIN 253
DB 334 KFRNLFRRFENLTDCRNHKSRYEIVSKKPPVLPPLIKDLTFLEHSGTFLVDGLVN 393
```

```
OY      254 FEKMMHRAARMLHCSSHNVPPLSRVSHLHEDSQVARISPCSEQSLSRBPAST 313
        |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db      394 IEKLHSVAKERTIKKRS-RLCLD-WASFNHL-----QT 428

OY      314 WAYVOQLKYIDNRELRLSRELE 337
        |||::|||::|||::|||::|||
Db      429 KAYROFGVIDNOLLFELSYKLE 452
```

```

RESULT 4
US-60-490-890-348
: Sequence 348 Application US/60490890
: GENERAL INFORMATION:
: APPLICANT: LI, Martha
: APPLICANT: Rupnow, Brent A.
: APPLICANT: Webster, Kevin R.
: APPLICANT: Jackson, Donald
: APPLICANT: Wong, Tai W.
: TITLE OF INVENTION: BIOMARKERS OF CYCLIN-DEPENDENT KINASE MODULATION
: FILE REFERENCE: D0310 PSP
: CURRENT APPLICATION NUMBER: US/60/490,890
: CURRENT FILING DATE: 2003-07-29
: NUMBER OF SEQ ID NOS: 2779
: SOFTWARE: Patentin version 3.2
: SEQ ID NO 348
: LENGTH: 1499
: TYPE: PRT
: ORGANISM: Homo sapiens
US-60-490-890-348

```

Query Match	23.3%	Score 409;	DB 7;	Length 1499;	
Best Local Similarity	30.9%	Pred. No. 5.9e-22;			
Matches 107; Conservative	57;	Mismatches 122;	Indels 60;	Gaps 5	

OY	35	VATSI:GLNERTLVNVPORVNEHLIPRPDQGLPTVGSAGEGLDYSAKDAOLQDTHDMSL	94
Db	675	LADRIQLSGRIYKLNKNMETETLCSDEBDADOLLRESOISLLOJSTVEVAQOLSRNRELE	734
OY	95	SIHOVELIHVYLGPOHLRDVTT-ANLEREPRRPNELQYVATELCLPVGPRAOALLRKE	1533
Db	735	NIEPTEYIDDLF---KLRSKTSICAMLKREEVYINQETFWVASIELRETQOLKRMKILKE	7911
OY	154	IKLAHLKEOKNLNLSFPMVMEGLSNATSRLHNTWERLPHKRYKLTSALERLIDPSPMNR	2133
Db	792	IKIALHCECKNFNSMFAIISGLNAPARLKTTEKLEPKYKLEQDLODLEDPBPRNMA	8511
OY	214	VYRLAL--AKLSPPVIFPMPLLLKDMTEIHGHNHTLVENLINFEXKMMARAARMLHCR	2711
Db	852	KYRWNLNSQNLQDPIILPFPVYIKDKLTFLEHGNDKSDVDLAVNEFKRLMIATKEIRHGRMA	9111
OY	272	SHNPVLSPLRLSR-----VSHVHEDS	292
Db	912	SYNMDPALMFRTRKKKKMSLGLSOGSGTATYALVDVAQOTGCHKKRVRRSSFLNAKKLYEDA	9711
OY	293	QVARISTCEQSLSTRSPASTWAYVOQLKVIDNQELSLRSLRELP	338
Db	972	QMAR-----KVQOYLSNLELMDDEESLQTLQDCEP	1002

```

: RESULT 5
: US-60-483-917-76
: Sequence 76, Application US/60483917
: GENERAL INFORMATION:
: APPLICANT: Gan, LI
: APPLICANT: Shiyak, David
: APPLICANT: Chin, Daniel J
: APPLICANT: von Schack, David
: APPLICANT: Uffer, Roman
: APPLICANT: Gonzalez-Zulueta, Mirella
: TITLE OF INVENTION: NUCLEIC ACIDS ASSOCIATED WITH NEURODEGENERATIVE DISORDERS
: FILE REFERENCE: 00208.0013.P2US00
: CURRENT APPLICATION NUMBER: US/60/483,917

```

```

; CURRENT FILING DATE: 2003-06-30
;
; NUMBER OF SEQ ID NOS: 78
;
; SOFTWARE: PatentIn version 3.2
;
; SEQ ID NO 76
;
; LENGTH: 557
;
; TYPE: PRT
;
; ORGANISM: Homo sapiens
US-60-483-917-76

```

Query Match	11.7%	Score 205.5	DB 7	Length 557
Best Local Similarity	24.6%	Pred. No. 8.5e-08		
Matches 69; Conservative	52	Mismatches 102	Indels 57	Gaps 9

QY	76	ISAKDLAQLDHDMSLNFISHQVEL-----HYVGLPQHLRVYTTANERFERREN	127
Db	49	VTPEEPASQITLMDLPEFAKIDPEELASCGSMKKESHSLAP-----NVAAETFRREN	99
QY	128	ELQYVAVATELCLCPRGPRQAOLRRFTKKAHLKEOKNLNSEFAVMEGLSNSAISRLAHT	187
Db	100	QVSPFVAVREILTAQTLKTAIAELISHFVKIAKKLELNLHSLMSVSVALQSAPIFRLTKT	159
QY	188	MEPLRHXYKRLYSALERLLDPSMNRHYVLAKLKS--PVYIPRMILLDMFTI---HEG	243
Db	160	WALLNRKKTKTTFEKIDLYLMSKEDNKKRFEYIRSLKMWSPISPLGLYLLDLDTLYIDSAYPA	219
QY	244	NHTLEVNLEINFEKMMMAFAARMLLHGRSHNVPLSPRSRVSHLHEDSQVARISTCSEQ	303
Db	220	SGSIMENQGRSNQMMNILR-----IADLQYV---SCSYD	250
QY	304	SLSTRSPASTW---AYVQQL-KVLDNQELSLRSLRELEP	338
Db	251	HLTTLPLHVQKYLSVRYIEELQKFEVDON--YKLSRIIRP	288

```

RESULT 6
US-60-487-610-1523
; Sequence 1523, Application US/60487610
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; APPLICANT: HUANG, Hongjin
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: LIVER FIBROSIS IN HEPATITIS C VIRUS-INJECTED
; TITLE OF INVENTION: METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: C1001469
; CURRENT APPLICATION NUMBER: US/60/487, 610
; CURRENT FILING DATE: 2003-07-17
; NUMBER OF SEQ ID NOS: 97101
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1523
; LENGTH: 1356
; TYPE: PRY
; ORGANISM: Homo sapiens
US-60-487-610-1523

```

Query Match	11.0%	Score 193;	DB 7;	Length 1356;
Best Local Similarity	27.9%	Pred No. 1.5e-06;		
Matches 60;	Conservative 46;	Mismatches 83;	Indels 26;	Gaps 9

```

OY      65  JTV-----GSAEELDVSAK--DLAOLDTHOMSLSNSHQVELINVLGRLOARDYT  115
      |||  :|  :|  :|  :|  :|  :|  :|  :|  :|  :|  :|  :|  :|  :|  :|  :|
Db      782  PTVEMHISRGHIEFDLTLTLPETIAQOLITLESOLDTAAORPSELVSQVWTKED--KEIN  840

OY      116  TANLEFRMRRELOVWVATELCLRPVGC--PRAOLLEKFIKLANLEKOKNLSFFAVM  173
      :|  :|  :|  :|  :|  :|  :|  :|  :|  :|  :|  :|  :|  :|  :|  :|
Db      841  SPNLLKMLRHTNTNLTF--EKCIYETENLEBRVAVNSRIFIELOFQELNFGVLEUV  898

OY      174  FGLNSAISRLAHTWERLPHKVKRLYSALERLLDPSWMHNRVYRLAKL--SPVPIRPM  230
      :|  :|  :|  :|  :|  :|  :|  :|  :|  :|  :|  :|  :|  :|  :|  :|
Db      899  SAMNSPVRYLDRHTEQIIPROKTI---LEAHNEISEDH--YKTKLAKLRSINPCVPRF  953

OY      231  PLLKDMTFINEGNTYV---ENLINFEMKRYMA  261
      :|  :|  :|  :|  :|  :|  :|  :|  :|  :|  :|  :|  :|  :|  :|
Db      954  GIYTLNLIKTEEGNEPVLLKRGKELINFSKRKKA  988

```

RESULT 7
US-60-487-610-1524
; Sequence 1524, Application US/60487610
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; APPLICANT: HUANG, Hongjin
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: LIVER FIBROSIS IN HEPATITIS C VIRUS-INFECTED SUBJECTS,
; FILE REFERENCE: C1001469
; CURRENT APPLICATION NUMBER: US/60/487,610
; CURRENT FILING DATE: 2003-07-17
; NUMBER OF SEQ ID NOS: 97101
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1524
; LENGTH: 1398
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-487-610-1524

Query Match 11.0%; Score 193; DB 7; Length 1398;
Best Local Similarity 27.9%; Pred. No. 1.5e-06;
Matches 60; Conservative 46; Mismatches 83; Indels 26; Gaps 9;
QY 65 PTV-----GSAEGLDVSAR--DLAQLTDHMSLFSNHQVELIHVGLGPQHLRDVT 115
DB 839 PTVWHISRPCHIEFDLTLHPRIARQLFLESDLYRAVQPSLVSQSVMTKED-KEIN 897
QY 116 TANLEFRMRRELOYWATELCCLPVG--PRAQLLKFKIKLAHKEQKNLSFFVAM 173
DB 898 SPNLKMRHRTTNLTLMF--ECIVETENLEERAVASRIIEILOVFQELNFGVLEVV 955
QY 174 FGLSSAISRLAHTWERLPHKVKRKLXSALERLLDPSNMHRVYRLAKL---SPVIVPEM 230
DB 956 SAMNSPYRLDHTQEQLPSROKTI---LEENHEISEH--YKTLAKLRISNPPCVFF 1010
QY 231 PLLKDMTFIHGHNHTLV---ENLINFEMKRMMA 261
DB 1011 GIYLTNMLKTEGNEPEVLKRKGKELINFSKRRKVA 1045

RESULT 8
US-60-487-610-2054
; Sequence 2054, Application US/60487610
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; APPLICANT: HUANG, Hongjin
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: LIVER FIBROSIS IN HEPATITIS C VIRUS-INFECTED SUBJECTS,
; FILE REFERENCE: C1001469
; CURRENT APPLICATION NUMBER: US/60/487,610
; CURRENT FILING DATE: 2003-07-17
; NUMBER OF SEQ ID NOS: 97101
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2054
; LENGTH: 1332
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-487-610-2054

Query Match 10.7%; Score 188.5; DB 7; Length 1332;
Best Local Similarity 24.8%; Pred. No. 3.1e-06;
Matches 60; Conservative 53; Mismatches 102; Indels 27; Gaps 9;
QY 29 QPDAAGVATSLGLENERLFVNVPQVEHLLIPPDQLGPTVGSAGLIDVSAR--DLAQLT 86
DB 740 QAQANGVSHNT-----TFESPPIEMHISKP-----GQFETFDLMTLHPRIARQLT 787
QY 87 DHMSLFRSHIQVELIHVGLGPQHLRDVTANLERFRNELQYWATELCCLPVG-- 144

DB 788 LLESDLYKVPQPSSELVGSWTKED-KEINSPNLKMRHRTTNLTLMF--ECIVAEENFE 844
QY 145 PRAQLRKFKIKLAHKEQKNLSFFAVMGFLSNSAISRLAHTWERLPHKVKRKLXSALER 204
DB 845 ERVAVLSRIITELQYFDQLNPNFVLEYSVANSVSVYRLDHTFEALQERRKI---LDE 901
QY 205 LIDPSWNH-RVYRLALAKLSPVPIPEMPLIKDMTFIHGHNHTLV---ENLINFEMKRM 259
DB 902 AVELSQDHFKRYLYLKLKSNINPCVFFGIYLTNMLKTEGNNDELKKKGKDLINFSRRK 961
QY 260 MA 261
DB 962 VA 963

RESULT 9
US-10-240-145A-86
; Sequence 86, Application US/10240145A
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 21272-048
; CURRENT APPLICATION NUMBER: US/10/240,145A
; CURRENT FILING DATE: 2002-09-27
; PRIOR APPLICATION NUMBER: 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 09/649,167
; PRIOR FILING DATE: 2000-08-23
; PRIOR APPLICATION NUMBER: 09/668,680
; PRIOR FILING DATE: 2000-09-22
; PRIOR APPLICATION NUMBER: 09/695,618
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 09/728,711
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: NOT YET ASSIGNED
; PRIOR FILING DATE: 2000-03-14
; NUMBER OF SEQ ID NOS: 172
; SOFTWARE: Custom
; SEQ ID NO 86
; LENGTH: 699
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-240-145A-86

Query Match 10.3%; Score 182; DB 6; Length 699;
Best Local Similarity 23.7%; Pred. No. 5e-06;
Matches 75; Conservative 52; Mismatches 118; Indels 72; Gaps 11;
QY 51 QEVHELPHPDQLGPTVGSAGLIDVSARKDLAGQLTDHMSLFSNHQVELIHVGLGPQ 110
DB 236 EEEGLMPOGROL-----LDF-SVDEVAEQTLTIDLELFSKVRLYEGLGSWSQRD 285
QY 111 LRDVTYAN--DERMRRENELOYWATELCCLP--VPGP-RAQLRKFKIKLAHKEQKNL 166
DB 286 RPGAAGSPYRAVAOENITVTCGLSVGLAPBLAOPRAORAKREKIRIAQORRELNF 345
QY 167 NSFPAVMGLSNSAISRLAHTWERLPHKVKRKLXSALERLLDPSNMHRVYRLAKL 219
DB 346 SSLRAILSALQSNITYLRLKSGAVSREPLSTFKLSQISDENNHLSREILHQEBATE 405
QY 220 -----AKLSPPVPEPMLIKDMTFIHGHNHTLV--NLINFEMKRM---MA 261
DB 406 GSOEDMTPGSLPSKPPGPVPYGLTFLDVLMDTALPMLLEGLDLNFERKREWEILA 465
QY 262 RAAMLHCHRNHNVPLSPLSRVSHLHEDSOVARISTCSROSISTSPASTMAYVQQLK 321
DB 466 RIQQLQRCOSYTUSPHPI---LAALHAONQLT-----EBQSY----- 501
QY 322 VIDNORELSRLELEP 338
DB 502 -----RLSRVIEP 509


```
RESULT 10
; Sequence 85, Application US/10240145A
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 21272-048
; CURRENT APPLICATION NUMBER: US/10/240,145A
; CURRENT FILING DATE: 2002-09-27
; PRIOR APPLICATION NUMBER: 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 09/649,167
; PRIOR FILING DATE: 2000-08-23
; PRIOR APPLICATION NUMBER: 09/668,680
; PRIOR FILING DATE: 2000-09-22
; PRIOR APPLICATION NUMBER: 09/695,618
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 09/728,711
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: NOT YET ASSIGNED
; PRIOR FILING DATE: 2000-03-14
; NUMBER OF SEQ ID NOS: 172
; SOFTWARE: Custom
; SEQ ID NO 85
; LENGTH: 715
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-240-145A-85

Query Match          10.3%; Score 182; DB 6; Length 715;
Best Local Similarity 23.7%; Pred. No. 5.1e-06;
Matches 75; Conservative 52; Mismatches 118; Indels 72; Gaps 11;

QY 51 QEVHELPHDQGPVGSNEGDLVSAKDLAGQLTDHMSLNSHNOVELHYVGLPQH 110
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 236 EEEEGMLPQGPOL-----LDF-SVEVAEQLTLDLELSKRYELGSLGVSQRD 285
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 111 LRDVTAN--LEFRMRFNELOYVATVTELCIP-VGCP-RAOLLRKFIAHLKEOKNL 166
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 286 RPAAGASPVYATVQAFNTVTCVGLSGVAGGLAPQAOAKLEKIRIAQACRELRNF 345
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 167 NSFFAVMGLSNSAISRLAHTWBLRPHKRYKLYSALERLDPMSNHRVYRLAL----- 219
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 346 SSLRAILSALQSNPIYRLKRSKMGAVSREPLSTFRKLQISQSDNNHLSSEILFOEATE 405
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 220 -----AKLSPPYIPFPMPLLKDMTFIHGNNHTLVE-NLINFEMR-----MMA 261
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 406 GSGQEDNTPGSLRSPKPPGVPYLTGFTLDTLVMLDTALPMLGDLINFERRKWEMLA 465
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 262 RAARMLHGRSHNPVPLSRVSHLHEDSOVARISTCSQSLSSTRSPASTWAVYQOLK 321
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 466 RIQQLORRCOSYLSLSPHPT---LALHAQNOULT-----EESQY----- 501
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 322 VIDNORELSRLSRELEP 338
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 502 -----RLSRVIER 509
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 11
; Sequence 1212, Application US/10296115
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL Nucleic Acids and Polypeptides
; FILE REFERENCE: 784PCT
; CURRENT APPLICATION NUMBER: US/10/296,115
; CURRENT FILING DATE: 2002-11-18
; PRIOR APPLICATION NUMBER: US09/488,725
; PRIOR FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: US09/552,317
; PRIOR FILING DATE: 2000-04-25
; NUMBER OF SEQ ID NOS: 1478
; SEQ ID NO 1212
```

```
; LENGTH: 821
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-296-115-1212

Query Match          8.4%; Score 147.5; DB 6; Length 821;
Best Local Similarity 21.7%; Pred. No. 0.0017;
Matches 66; Conservative 49; Mismatches 118; Indels 71; Gaps 10;

QY 78 AKDL-AGQLTDHMSLFPNSHNOVELHYVGLG-----PQHLRDVTYTNLERPMRRF 126
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 284 SEDLVAEQLTYMDQLEFKV---VPHHCIGCTMSRRDKENKHNLAFTIATIS---QP 335
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 127 NELQYVATPELCSPVPG-----BRAOLLRKFIAHLKEOKNLNSFFAVMGLSN 179
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 336 NTL-----TKCVSTLDGKELTKQORAKITEWIMINAECHRLKNFSSRLATVSLQSN 350
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 180 AISRLAHTWBLRPHKRYKLYSALERLDPMSNHRVYRLALAK----- 221
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 391 SIYRLAKTMAVAPRDRMLMEELSDIFSDHNNHLSRELLMKRGTSFANLDSSVKNOK 450
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 222 -----LSPPVLPFPMPLLKDMTFIHGNNHTLVE-NLINFEMR-----MMA 263
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 451 RTQRRQLQDKMGVMQGYVYLGFTLDTLMTALQDYIEGLINFEKRRRFEVYAOI 510
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 264 ARMLHGRSHNPV-----LSPLRSVSHLHEDSOVARISTCSQSLSSTRSPASTWAVYQ 319
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 511 KLIQSACNSYCMTPDQKFIOMFOROOLLTEESIALCEIAAADASTTSPKPKSMVKR 570
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 320 LKVI 323
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 571 LNL 574
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 12
; Sequence 1754, Application US/60487610
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: LIVER FIBROSIS IN HEPATITIS C VIRUS-INFECTED SUBJECTS,
; FILE REFERENCE: CL001469
; CURRENT APPLICATION NUMBER: US/60/487,610
; CURRENT FILING DATE: 2003-07-17
; NUMBER OF SEQ ID NOS: 97101
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1754
; LENGTH: 814
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-487-610-1754

Query Match          6.1%; Score 108; DB 7; Length 814;
Best Local Similarity 21.9%; Pred. No. 1.1;
Matches 79; Conservative 50; Mismatches 120; Indels 112; Gaps 16;

QY 31 DARGVATSLGNE--RLFVNVPOEVHELPHDQGLGTVGSAGGLDLYSAKDLAQLTDH 88
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 342 DRGQVITMQALSEDRLLMEAMDGRPYVNSNMDSQSESTA-----QLDSI 388
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 89 DMSLFNS-INOVELINHYVGLPQHLRDVTYTNLERPMRFELOYVNA-----T 135
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 389 GFSIIRKCHAVE-----TRGINEQGLYRIGVNSRVOKLLSLVMDPTASSTET 438
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 136 ELG-----LCVPVGP--RAOLLRKFIAHLKEOKNLNSFFAVMGLSN 178
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 439 DICAEMIKITTSALKTYLMLRGPMLMYQFQSFIR-AKLENQ----- 483
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 179 SAISRLAHTWBLRPHKRYKLYSAL-ERLDPMSNHRVYRLALAKLSPPVLPFPMPLLK-- 235
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 484 SRVSEIHSIVHRRLPEKRRQMLQLMNLHANVANNNHKONLMTYVNLG---VVGSPITLLRPQ 540
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
```

QY 236 -----DMTFIEGHNLTLENLINFEXMMARAARMLHHCRRSHNPVPLSLRSVSH 287
DB 541 EETVAIMDKFQNTVIELLEN-----HEKIFNTPDMLTNALH 582
QY 288 L-HESQVARICTSEOSIST-RSPASTWAYVQOLKVID-----NORELSRLSRELE 337
DB 583 LSRKSSDPSKPPSCSESRPLTFHTVOSTEKORNSIINSSLESVSSNPNSTILNSSSLQ 642
QY 338 P 338
DB 643 P 643
RESULT 13
US-10-425-114A-46385
; Sequence 46385, Application US/10425114A
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovacic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(5313)B
; CURRENT APPLICATION NUMBER: US/10/425,114A
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 46385
; LENGTH: 437
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: 700100978_FLI.pep
US-10-425-114A-46385
Query Match 5.8%; Score 101.5; DB 6; Length 437;
Best Local Similarity 21.7%; Pred. No. 1.8;
Matches 72; Conservative 35; Mismatches 98; Indels 127; Gaps 14;
QY 2 AALAGDEGWTGQGVLYVKNSSGDAIGLQPDARGVATSLGRLNRLVYVPOEVH----- 54
DB 148 AAVADERVYQGEV-----AGDDG-----HAGVV-----LLVHPRRELHVGVGVD 190
QY 55 -----ELIPH-----PQOLGPTGSA 70
DB 191 VHQGVHDLVHVRVYLGAAHATRAGVYVDEAGHLALLDVGGTLVALLPDELG----- 244
QY 71 BGLDLYSAKDLAQLTDHMSLENSIHOVELIHVYLGPOHLRDVTANLEREMRRF--- 126
DB 245 -GLSGVALQLAG-----AHHGGAHLGEPDLA-LERESLAFASPD 284
QY 127 -----NELQYVATELCIPVPGRRQQLRKFKILAAHLKEQKNLSFPAVWGLSNSAIS 182
DB 285 TOYBSNLDVWQCHEV-LGEVDEDELVEALGYEALIHVEVVRLLALADVYVAGHGDVG 343
QY 183 RLATHTWERLPHKVRKLYSALERLDPSSNNHRYRLALAKLSPPVLPFMPILLKDKDTFIH- 241
DB 344 AVPH-----RRLLEVVDALAPGVDLLVVRDVGQALPRV-----LVRLRLVHL 388
QY 242 EGNHTLVENLINFEXMMARAARMLHHCRRSH 273
DB 389 LGSHTL-----EPAAVAGALAMVGHGRGH 412
RESULT 14
PCT-US03-22819-2
; Sequence 2, Application PC/TUS0322819
; GENERAL INFORMATION:
; APPLICANT: PELICER, Angel
; APPLICANT: LEONARDI, Peter

; APPLICANT: INGHIRAI, Giorgio
; TITLE OF INVENTION: HUMAN RGR ONCOGENE AND TRUNCATED TRANSCRIPTS THEREOF DETECTED
; TITLE OF INVENTION: T CELL MALIGNANCIES, ANTIBODIES TO THE ENCODED POLYPEPTIDES A
; TITLE OF INVENTION: METHODS OF USE
; FILE REFERENCE: PELICER-1A PCT
; CURRENT APPLICATION NUMBER: PCT/US03/22819
; CURRENT FILING DATE: 2003-07-23
; PRIOR APPLICATION NUMBER: 60/397,873
; PRIOR FILING DATE: 2002-07-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 2
; LENGTH: 473
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US03-22819-2
Query Match 5.6%; Score 99; DB 1; Length 473;
Best Local Similarity 20.7%; Pred. No. 3;
Matches 69; Conservative 52; Mismatches 122; Indels 90; Gaps 14;
QY 47 VVNPQEVHELPHPPQDGPVGSAGLDLVSAKD-----LAGQLTDHMSLEFNSIHQ 98
DB 187 VLEPQSA-----PSSCPGSGSVKNQPSSELPMTTPPRLIAQLTLMDELFKRYVL 240
QY 99 VELIHVYLGPOHLR--DVTANLEREMRRFNEQYVATELCPCVPG-----PRAQL 149
DB 241 HECLICWGGHLLKGNENHMAPTVRATIAHFNLTCITTS-CL-----GDHSMRADRRARY 295
QY 150 LRKFKILAAHLKEQKNLSFPAVWGLSNSAISRLAHWELPHK-----VKRLYSAL 202
DB 296 VEHWIKVARCECLSNFSSVHVYALCSNPGQLHKTWAVSSSKMELKELCKDTAV 355
QY 203 ER-LIDPSSMNRV-----YRLALAKLSPPVLPFMPILLKDKDTFIH-----EGNH 245
DB 356 KRDLIKRGSFRKVAQENRPOVQRLRQKKGVYPLFDLTELQRLDSAIPODLQNT 415
QY 246 TLVENLINFEXMMARAARMLHHCRRSHNPVPLSLRSVSHLEDQVARICTSEQSL 305
DB 416 NKRSKEVHVLDEMQLQVAAAMYRLR-----PLEKFVYF----- 450
QY 306 STRPASTWAYVQOLKVIDNORELSRLSRELEP 338
DB 451 -TR-----MEQL-----SPKESYKLSQLEP 470
RESULT 15
US-10-425-114A-40595
; Sequence 40595, Application US/10425114A
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovacic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(5313)B
; CURRENT APPLICATION NUMBER: US/10/425,114A
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 40595
; LENGTH: 438
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB189-024-D7_FLI.pep
US-10-425-114A-40595
Query Match 5.4%; Score 94.5; DB 6; Length 438;
Best Local Similarity 21.4%; Pred. No. 5.8;
Matches 71; Conservative 35; Mismatches 99; Indels 127; Gaps 14;

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 7, 2003, 07:11:44 : Search time 29 seconds
(without alignments)
493.140 Million cell updates/sec

Title: US-09-856-679-2

Perfect score: 1759

Sequence: 1 MALAODEGWTGQVLKVN.....QLKVIDNORELSRLSELP 338

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued_Patents_AA:*

1: /cgn2_6/ptodata/1/laa/5A_COMB.pep:*

2: /cgn2_6/ptodata/1/laa/5B_COMB.pep:*

3: /cgn2_6/ptodata/1/laa/6A_COMB.pep:*

4: /cgn2_6/ptodata/1/laa/6B_COMB.pep:*

5: /cgn2_6/ptodata/1/laa/PCTUS_COMB.pep:*

6: /cgn2_6/ptodata/1/laa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	213	12.1	557	US-09-568-816A-2	Sequence 2, Appl1
2	208.5	11.9	911	US-09-356-952-6	Sequence 6, Appl1
3	206	11.7	1244	US-09-356-952-7	Sequence 7, Appl1
4	202.5	11.5	402	US-08-290-731C-14	Sequence 14, Appl1
5	199	11.3	489	US-08-318-831-4	Sequence 4, Appl1
6	199	11.3	666	US-08-318-831-3	Sequence 3, Appl1
7	199	11.3	814	US-08-318-831-2	Sequence 2, Appl1
8	194	11.0	364	US-08-318-831-6	Sequence 6, Appl1
9	193	11.0	423	US-08-290-731C-10	Sequence 10, Appl1
10	193	11.0	684	US-09-765-298A-18	Sequence 18, Appl1
11	193	11.0	1319	US-08-290-731C-2	Sequence 2, Appl1
12	193	11.0	1333	US-09-356-952-2	Sequence 2, Appl1
13	193	11.0	1336	US-08-290-731C-6	Sequence 6, Appl1
14	189.5	10.8	652	US-08-318-831-8	Sequence 8, Appl1
15	181	10.3	418	US-08-290-731C-13	Sequence 13, Appl1
16	181	10.3	1048	US-09-356-952-5	Sequence 5, Appl1
17	178	10.1	426	US-08-290-731C-12	Sequence 12, Appl1
18	178	10.1	430	US-08-290-731C-5	Sequence 5, Appl1
19	178	10.1	1572	US-08-290-731C-11	Sequence 11, Appl1
20	178	10.1	1589	US-09-356-952-4	Sequence 4, Appl1
21	178	10.1	1596	US-09-356-952-3	Sequence 3, Appl1
22	177.5	10.1	423	US-08-290-731C-11	Sequence 11, Appl1
23	177.5	10.1	1297	US-08-290-731C-4	Sequence 4, Appl1
24	164.5	9.4	362	US-08-290-731C-15	Sequence 15, Appl1
25	153.5	8.7	852	US-08-408-519-5	Sequence 5, Appl1
26	153.5	8.7	852	PCT-US95-03552-5	Sequence 5, Appl1
27	143	8.1	768	US-08-408-519-2	Sequence 2, Appl1

28	143	8.1	768	5	PCT-US95-03552-2	Sequence 2, Appl1
29	127.5	7.2	688	4	US-09-367-206-20	Sequence 20, Appl1
30	127.5	7.2	703	4	US-09-367-206-5	Sequence 5, Appl1
31	121.5	6.9	501	4	US-09-367-206-3	Sequence 3, Appl1
32	103.5	5.9	1504	4	US-09-364-206-2	Sequence 2, Appl1
33	96	5.5	337	4	US-09-252-911A-257198	Sequence 257198, A
34	95.5	5.4	452	4	US-09-252-911A-26867	Sequence 26867, A
35	93	5.3	462	3	US-09-055-113-4	Sequence 4, Appl1
36	91	5.2	486	3	US-08-348-518C-2	Sequence 2, Appl1
37	91	5.2	515	3	US-08-930-966A-12	Sequence 12, Appl1
38	88.5	5.0	473	1	US-08-103-739B-2	Sequence 2, Appl1
39	88.5	5.0	473	2	US-08-474-404-2	Sequence 2, Appl1
40	88.5	5.0	473	2	US-08-485-845-2	Sequence 2, Appl1
41	88.5	5.0	473	2	US-08-482-714-2	Sequence 2, Appl1
42	88.5	5.0	473	3	US-09-211-416-2	Sequence 2, Appl1
43	88.5	5.0	473	3	US-09-059-958-2	Sequence 2, Appl1
44	88	5.0	639	4	US-09-252-911A-25089	Sequence 25089, A
45	84.5	4.8	382	3	US-08-660-645A-7	Sequence 7, Appl1

ALIGNMENTS

RESULT 1
US-09-568-816A-2
: Sequence 2, Application US/09568816A
: Patent No. 6440699
: GENERAL INFORMATION:
: APPLICANT: Tavrilgian, Sean V.
: APPLICANT: Swedlund, Brad
: APPLICANT: Simard, Jacques
: APPLICANT: Rommens, Johanna M.
: APPLICANT: Myriad Genetics, Inc.
: APPLICANT: Hospital for Sick Children
: TITLE OR INVENTION: CA7 CG04 Gene
: FILE REFERENCE: 2318-237-11
: CURRENT APPLICATION NUMBER: US/09/568, 816A
: CURRENT FILING DATE: 2000-05-11
: PRIOR APPLICATION NUMBER: US 60/134, 209
: PRIOR FILING DATE: 1999-05-14
: NUMBER OF SEQ ID NOS: 21
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 2
: LENGTH: 557
: TYPE: PRT
: ORGANISM: Homo sapiens
: US-09-568-816A-2

Query Match 12.1%; Score 213; DB 4; Length 557;

Best Local Similarity 27.1%; Pred. No. 4.2e-14;
Matches 70; Conservative 45; Mismatches 113; Indels 30; Gaps 7;

QY	76	VSADLADGLTIDHMSLFNSIHVELI-----HYLDPOHLRDVTTANLEFRMRFN 127	
DB	48	VTEPEVYGOITLDVDPFKAIODELSSCGMKNKEKYSAP-----NAVAFTFRFN 98	
QY	128	ELOYWVATELCPCVPQOLKRFIKLAHLKEOKNLNSFPVFMGLNSATSRLAHT 187	
DB	99	HASWVYRELHAQOTLIRAEVLSHYTKTKKLYELNNLAHVAVSGLSAPFRLTKT 158	
QY	188	WERLPHKVRKLYSALERLDPSSNHRVRLAKLS--PVYIPFMPLLKDMTFIEGHNHT 246	
DB	159	WALLSRDKTTFPEKLEVMSKEDNYKRLROYISSLKMTPCIPYLGIVLSUTYIDSAVPS 218	
QY	247	LVENLIFEKRMARARALHHCRS--HNVPVLSP---LRSRVSLHE-----DSOVA 295	
DB	219	TGSILEHORSNNMNNLRISDLOOCEVDIPMLPHVQRYLNSVOYIEELQKVEDDNY 278	
QY	296	RISCSQSLST--RSPAS 312	
DB	279	KLSLKIPGTSTPSAAS 296	

Query Match	11.0%;	Score 193;	DB 2;	Length 1319;
Best Local Similarity	27.9%;	Pred. No. 2,2e-11;		
Matches	60;	Conservative	46;	Mismatches 83; Indels 26; Gaps 9
OY	65	PTV-----GSAEGLDLSAK--DLAGQLTDHMSLFSNTHQVELIHVYLGPQHLRDVT	115	
Db	759	PTVMHISRPHIEFDLTLTPRIEIAQLTLESDLRVAOPSELVSVMTKED--KEIN	817	
OY	116	TANERFERRENELOYWATELCGLPVBG--PRAOLKRFKTLAHNKEOKLSNFPVAM	173	
Db	818	SPNLKMTIRHTNLTLM--EKCIYETENLEERVAVSRIIFILOFQELNNFNGVLEV	875	
OY	174	FGLSNATSRLAHTWERLPHKVRKLYSALERLLDPSWMHRYVRLALAKL--SPVIFPM	230	
		:::	:::	:::
Db	876	SAMSSPVYRLDHTFEQIPRSROKTI---LEEANHELSEDH--YKKYLAKLRSINPCVPDFE	930	
OY	231	PLLLKDMFIEHGNGHTLV-----ENINIEKKRMMA	261	
		::	::	::
Db	931	GIYLTNIIKTEGNEPVLRRRGKELINSSKRRVA	965	

```

US-09-356-952-2
RESULT 12
US-09-356-952-2
: Sequence 2, Application US/09356952
: Patent No. 6117663
: GENERAL INFORMATION:
: APPLICANT: Borlack-Sjodin, Ann
: APPLICANT: Margalit, S. M.
: APPLICANT: Bor-Sogit, Dafna
: APPLICANT: Cole, Philip
: APPLICANT: Kuriyan, John
: TITLE OF INVENTION: A CRYSTAL OF A RAS-SOS COMPLEX AND METHODS OF USE
: TITLE OF INVENTION: THEREOF
: FILE REFERENCE: 600-1-228N
: CURRENT APPLICATION NUMBER: US/09/356,952
: CURRENT FILING DATE: 1999-07-19
: EARLIER APPLICATION NUMBER: 60/093,631
: EARLIER FILING DATE: 1998-07-21
: NUMBER OF SEQ ID NOS: 14
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 2
: LENGTH: 1333
: TYPE: PR1
: ORGANISM: Homo sapiens
US-09-356-952-2

```

	Query Match	11.0%;	Score 193;	DB 3;	Length 1333;
	Best Local Similarity	27.9%;	Pred. No. 2,3e-11;		
	Matches	60;	Conservative	46;	Mismatches 83; Indels 26; Gaps 9
OY	65 PTV-----GSAEGDLVSAK--DLAGQLTDHDSLFNSIHQVELHYVLGPOLRDVT	115			
Db	759 PTVMHSIRPCHIEFTFDLTLHPRIEALQTLLSESDLRAPOPSLSVMTKEO-KETIN	817			
OY	116 TANERIFRRNELEYWATELCSPRG--PRAGLKRFKLANTKEOKNLSEFPAVM	173			
Db	818 SPNLIKTRHTNTLTLP--EKCIYETENLEERAUVAVSRIIEILOVFPELNNGVELEV	875			
OY	174 FGLSNSAISLAHTWERLPHKVRKLYSALERLDPSWMHRVYLALAK--SPVIDPFM	230			
Db	876 SAMSSPYRDRDHFEQIPSRQKKI---LEEAHELSEDH--YKKYIAKTIRSINPCVPBE	930			
OY	231 PLLDKDMTFINEGHNTLV----ENINTEKKRMMA	261			
Db	931 GIYTLNIILKTEEGNPVELKRKGKELINSKRRKA	965			

RESULT 13
US-08-290-731C-6
; Sequence 6, Application US/08290731C
; Patent No. 5843646
; GENERAL INFORMATION:

```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000

```

TITLE OF INVENTION: PREPARATION AND UTILIZATION
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: Rhone-Poulenc Rorer Inc.
STREET: 500 Arcola Road, 3C43
CITY: Collegeville
STATE: PA
COUNTRY: USA
ZIP: 19426
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: Macintosh
OPERATING SYSTEM: System 7.1
SOFTWARE: Word 5.1 (EPO PatentIn)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/318, 831
FILING DATE: 19 October 1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: FR92/04827
FILING DATE: 21-APR-1992
ATTORNEY/AGENT INFORMATION:
NAME: Smith, Julie K.
REGISTRATION NUMBER: P-38,619
REFERENCE/DOCKET NUMBER: ST92033-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (610) 454-3839
TELEFAX: (610) 454-3808
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 652 amino acids
TYPE: amino acid
TOPOLOGY: linear
US-08-318-831-8

Query Match 10.8%; Score 189.5; DB 1; Length 652;
Best Local Similarity 24.8%; Pred. No. 1.8e-11;
Matches 60; Conservative 53; Mismatches 102; Indels 27; Gaps 9;

QY 29 QDARQVATSLGLENLFPVNVQEVHELIRHPDQCPYVGSAGLDLVSK--DLGQLT 86
DB 97 QAVANVSHNI-----TFESPPIPIEMHISKP-----GQETFDMLHPLEIAROLT 144
QY 87 DHDMSLFNSIHQVELIHYVGLPOHLRDVTTANLERMRRENELQYVWATELCSPVGC-- 144
DB 145 LLESIDYRKVQPSVETGVSMTKED-KEINSPNLKMRIRHTNLTLMF--EKCIYEENNE 201
QY 145 PRAQLRKFKIKLAHLKEQKNLNSFPAVNFGLSNSAISRLAHTWERLPHKVRKLYSALER 204
DB 202 ERVAVALSRILEILOVRDLNFNGLVETISAVNSVSVRLDHTFEALQEKRRKI--LDE 258
QY 205 LLDPSNNH-RVYRLALAKLSPYIPMPPLLDKMTIHGNGHTLV----ENLINFEMRM 259
DB 259 AVELSDHFKKYLVLKLSINPCVVPFGIYLTNLTKEEGNDFLKKKGKDLINFSKRRK 318
QY 260 MA 261
DB 319 VA 320

RESULT 15
US-08-290-731C-13
Sequence 13. Application US/08290731C
Patent No. 5843646
GENERAL INFORMATION:
APPLICANT: BOWTELL, David Douglas Lawrence
TITLE OF INVENTION: DNA MOLECULES ENCODING MURINE
TITLE OF INVENTION: SON OF SEVENTEENLESS (MSOS) GENE,
TITLE OF INVENTION: AND MSOS POLYPEPTIDES
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESSEE: SUGHRUE, MION, ZINN, MACPEAK & SEAS
STREET: 2100 PENNSYLVANIA AVENUE, N.W.

CITY: WASHINGTON
STATE: D.C.
COUNTRY: USA
ZIP: 20037
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/290, 731C
FILING DATE: 17-OCT-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/AU93/00068
FILING DATE: 17-FEB-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PL0921/92
FILING DATE: 17-FEB-1992
ATTORNEY/AGENT INFORMATION:
NAME: KIT, Gordon
REGISTRATION NUMBER: 30,764
REFERENCE/DOCKET NUMBER: Q-36066
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 293-7060
TELEFAX: (202) 293-7860
INFORMATION FOR SEQ ID NO: 13:
SEQUENCE CHARACTERISTICS:
LENGTH: 418 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-290-731C-13

Query Match 10.3%; Score 181; DB 2; Length 418;
Best Local Similarity 23.2%; Pred. No. 7.4e-11;
Matches 73; Conservative 57; Mismatches 144; Indels 40; Gaps 8;

QY 14 QVLYKNSAGDAIGLDPDARGVATSLGLENLFPVNVQEVHELIRHPDQ-----LGPT 66
DB 107 QVAKENINGSVELLE-----VNCKFLGNIOETATPMKTLDDOICODHYSGLT 156
QY 67 VGSAGGLDVSADKDLGOLTDHDMSLFNSIHQVELIHYVGLPOHLRDV--TTANLERMR 125
DB 157 YSTTESILAVDPVLFATQLTILEHEITYCETTTDCLOKIKKNTYSYGASPOLNEFISF 216
QY 126 FNELOYVATELCICVPGPRAQLRKFKIKLAHLKEQKNLNSFPAVNFGLSNSAISRLA 185
DB 217 ANKLTNPISSVYKKAADSKSRRAKLSHFIFIAEYCKRFNFSMTDIIISALYSSPIYRLE 276
QY 186 HWERLPHKVRKLYSALERLDDPSMNRVYRLALAKL-SPYIPMPPLLDKMTIHGNG 244
DB 277 KTWQAAVIPQTRDLOSINKLMDPKKPFINRNELKSLHSAFCVVPFGVYLSLDTFTDSGN 336
QY 245 -----HTL-----VENLINFEK-MRMMAAABMLHHCASHNPVP-----LSPLRS 283
DB 337 PDYLVLEHGLKGVHDEKKTINENKRSRLDILQELITFKKTHYDFTKDRIVICISNSLE 356
QY 284 RVSHLEDSSQVARI 297
DB 397 NIPHEKQYQSLI 410

Search completed: October 7, 2003, 07:16:48
Job time : 31 secs

THIS PAGE BLANK (USPTO)


```

FT Modified-site /note="Potential phosphorylation site"
FT 299 /note="Potential phosphorylation site"
FT Modified-site 306
FT /note="Potential phosphorylation site"
XX
XX MO200031263-A2.
XX
XX
XX 02-JUN-2000.
XX
XX 23-NOV-1999; 99WO-US28013.
XX
XX 23-NOV-1998; 98US-0109592.
XX 04-FEB-1999; 99US-0118610.
XX 06-APR-1999; 99US-0127990.
XX
XX (INCY-) INCYTE PHARM INC.
XX
XX Hillman JL, Tang YT, Bandman O, Lal P, Yue H, Lu DM, Baughn MR;
XX Yang J, Azimzai Y;
XX
XX WPI; 2000-400073/34.
XX N-PSDB; AAA49172.
XX
XX Human GTPase associated proteins, polynucleotides, and antibodies,
XX useful for diagnosing, preventing and treating various diseases such as
XX atherosclerosis, cancer, acquired immune deficiency syndrome (AIDS),
XX asthma, and autoimmune diseases -
XX
XX Claim 1; Page 83-84; 144pp; English.
XX
XX Human cDNA libraries from various tissues were screened for GTPase
XX associated proteins (GTPAP). The present sequence is human
XX GTPAP-2 protein. This sequence was derived from a cDNA library of the
XX brain meningioma tissue from a 35 year old female. This protein is
XX expressed in reproductive, nervous and cardiovascular tissue. The
XX GTPAP proteins may be used to define agonists and antagonists of GTPAP
XX activity and to generate antibodies to GTPAP. This means the GTPAP
XX proteins may be useful for treatment or prevention of diseases
XX associated with GTPAP such as cell proliferation disorders, autoimmune
XX disorders, inflammatory disorders, immune system disorders, cancer,
XX AIDS, asthma, atherosclerosis, arthritis, systemic lupus erythematosus
XX and psoriasis.
XX
XX Sequence 338 AA:
SQ
Query Match 100.0%; Score 1759; DB 21; Length 338;
Best Local Similarity 100.0%; Pred. No. 1,8e-169;
Matches 338; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

RESULT 2
ID AAY70965
XX AAY70965 standard; protein; 881 AA.
XX
XX AAY70965;
XX
XX 09-AUG-2000 (first entry)
XX
XX Human Ras signalling pathway associated protein CAMP-GEFI.
XX
XX Human; Ras signalling pathway; CAMP-GEFI; cyclic adenosine monophosphate;
XX GEF; guanine nucleotide exchange factor; Rap1A; diagnosis; treatment;
XX CAMP-GEF-associated disorder; drug; transgenic animal model;
XX Ras-associated cancer; protein therapy.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX Binding-site 231..300
XX /label=CAMP-binding_domain
XX Region 278..282
XX /note="conserved motif in the CAMP binding pocket;
XX amino acid at position 280 determines binding
XX specificity for CAMP/CGMP"
XX Region 616..639
XX /label=SCR1
XX /note="Structurally conserved region which is
XX highly homologous to Ras-superfamily GEFs"
XX Region 689..731
XX /label=SCR2
XX /note="Structurally conserved region which is
XX highly homologous to Ras-superfamily GEFs"
XX Region 768..789
XX /label=SCR3
XX /note="Structurally conserved region which is
XX highly homologous to Ras-superfamily GEFs"
XX
XX MO200024768-A2.
XX
XX 04-MAY-2000.
XX
XX 22-OCT-1999; 99WO-US24826.
XX
XX 23-OCT-1998; 98US-0105507.
XX 16-NOV-1998; 98US-0108685.
XX
XX (MAST ) MASSACHUSETTS INST TECHNOLOGY.
XX
XX Kawasaki H, Graybiel A, Housman D;
XX WPI; 2000-350690/30.
XX N-PSDB; AAD00315.
XX
XX Isolated nucleic acid comprises nucleotide sequence encoding protein
XX selected from normal or mutant CalDAG-guanine nucleotide exchange
XX factor
XX
XX Claim 64; Page 113-115; 128pp; English.
XX
XX The present sequence is a human cyclic adenosine 3', 5'
XX monophosphate-guanine nucleotide exchange factor (CAMP-GEFI) which
XX has CAMP binding domain and Ras superfamily GEF
XX domains. It has substrate specificity for Rap1A and is differentially
XX distributed in brain and various other tissues. It selectively activates
XX the Ras superfamily small G protein and Rap1A, functions as a negative
XX regulator of Ras and directly couples the CAMP signal transduction
XX system to Ras superfamily cascades. The present sequence
XX is used for diagnosis and treatment (by protein therapy) of
XX CAMP-GEF-associated disorders. It is also useful for developing drugs
XX and producing cell-lines or transgenic animal models for Ras-associated
XX cancers.

```

XX Sequence 881 AA:
 SQ Query Match 99.4%; Score 1748; DB 21; Length 881;
 Best Local Similarity 99.1%; Pred. No. 9.4e-168;
 Matches 335; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 MAALAEDEGWTGKGVLYKVNAGDAIGLPDARGVATSLGNERLFPVNPQEVHELPHR 60
 DB 544 MAALAEDEGWTGKGVLYKVNAGDAIGLPDARGVATSLGNERLFPVNPQEVHELPHR 603
 QY 61 DQLGPTVGSAGELDLVSAKDLAQGLTDHWSLFSNIHQVELIHVLCRQHLRDVTYANLE 120
 DB 604 DQLGPTVGSAGELDLVSAKDLAQGLTDHWSLFSNIHQVELIHVLCRQHLRDVTYANLE 663
 QY 121 RFRMRFNELOYWATELCLCPVPGPRAQLLRFKTKLAHLKEOKNLSFPAVMFGLSNSA 180
 DB 664 RFRMRFNELOYWATELCLCPVPGPRAQLLRFKTKLAHLKEOKNLSFPAVMFGLSNSP 723
 QY 181 ISRLAHTWERLPHKVRKLYSALERLDPSPNNHRYRLALAKLSPPIVFPMLLLKDMTFI 240
 DB 724 ISRLAHTWERLPHKVRKLYSALERLDPSPNNHRYRLALAKLSPPIVFPMLLLKDMTFI 783
 QY 241 HEGNHTLVENLINFEEKRMMAARAAMLHCRSHNPVPLSPRSRVSHLHEDSOVARISTC 300
 DB 784 HEGNHTLVENLINFEEKRMMAARAAMLHCRSHNPVPLSPRSRVSHLHEDSOVARISTC 843
 QY 301 SEOSTSRPASTWAVYQOLKVIDNORELSRLSRELEP 338
 DB 844 SEOSTSRPASTWAVYQOLKVIDNORELSRLSRELEP 881

RESULT 3
 AAY70964 ID AAY70964 standard; Protein: 884 AA.
 AC AAY70964;
 DT 09-AUG-2000 (first entry)
 XX
 DE Rat Ras signalling pathway associated protein CAMP-GEFI.
 XX
 KM Rat; Ras signalling pathway; CAMP-GEFI; cyclic adenosine monophosphate;
 KM GEF; guanine nucleotide exchange factor; Rap1; diagnosis; treatment;
 KM CAMP-GEF-associated disorder: drug; transgenic animal model;
 KM Ras-associated cancer; protein therapy.
 XX
 OS Rattus norvegicus.
 XX
 FH Key Location/Qualifiers
 FH Blinding-site 231..300
 FT /label= CAMP-binding-domain
 FT Region 278..282
 FT /note= "conserved motif in the CAMP binding pocket;
 FT amino acid at position 280 determines binding
 FT specificity for CAMP/cGMP"
 FT Region 619..642
 FT /label= SCRI
 FT /note= "Structurally conserved region which is
 FT highly homologous to Ras-superfamily GEFs"
 FT Region 692..734
 FT /label= SCR2
 FT /note= "Structurally conserved region which is
 FT highly homologous to Ras-superfamily GEFs"
 FT Region 771..792
 FT /label= SCRI
 FT /note= "Structurally conserved region which is
 FT highly homologous to Ras-superfamily GEFs"
 PN WO200024768-A2.
 XX 04-MAY-2000.
 XX

PF 22-OCT-1999; 99WO-US24826.
 XX
 PR 23-OCT-1998; 98US-0105507.
 PR 16-NOV-1998; 98US-0108685.
 XX
 PA (MAST) MASSACHUSETTS INST TECHNOLOGY.
 XX
 PI Kawasaki H, Graybiel A, Housman D;
 XX
 DR WPI: 2000-350690/30.
 DR N-PSDB: AAD00314.
 XX
 PT Isolated nucleic acid comprises nucleotide sequence encoding protein
 PT selected from normal or mutant CalDAG-guanine nucleotide exchange
 PT factor -
 XX
 PS Claim 64; Page 107-109; 128pp; English.
 XX
 CC The present sequence is a cyclic adenosine 3', 5'
 CC monophosphate-guanine nucleotide exchange factor (CAMP-GEFI) from rat.
 CC The CAMP-GEFI has CAMP binding domain and Ras superfamily GEF
 CC domains. It has substrate specificity for Rap1 and is differentially
 CC distributed in brain and various other tissues. It selectively activates
 CC the Ras superfamily small G protein and Rap1, functions as a negative
 CC regulator of Ras and directly couples the CAMP signal transduction
 CC system to Ras superfamily cascades. The present sequence
 CC is used for diagnosis and treatment (protein therapy) of
 CC CAMP-GEF-associated disorders. It is also useful for developing drugs
 CC and producing cell-lines or transgenic animal models for Ras-associated
 CC cancers.
 CC
 XX
 SQ Sequence 884 AA;
 SQ
 XX
 Query Match 94.6%; Score 1664; DB 21; Length 884;
 Best Local Similarity 93.5%; Pred. No. 3.1e-159;
 Matches 316; Conservative 11; Mismatches 11; Indels 0; Gaps 0;

QY 1 MAALAEDEGWTGKGVLYKVNAGDAIGLPDARGVATSLGNERLFPVNPQEVHELPHR 60
 DB 547 MAALAEDEGWTGKGVLYKVNAGDAIGLPDARGVATSLGNERLFPVNPQEVHELPHR 606
 QY 61 DQLGPTVGSAGELDLVSAKDLAQGLTDHWSLFSNIHQVELIHVLCRQHLRDVTYANLE 120
 DB 607 DQLGPTVGSAGELDLVSAKDLAQGLTDHWSLFSNIHQVELIHVLCRQHLRDVTYANLE 666
 QY 121 RFRMRFNELOYWATELCLCPVPGPRAQLLRFKTKLAHLKEOKNLSFPAVMFGLSNSA 180
 DB 667 RFRMRFNELOYWATELCLCPVPGPRAQLLRFKTKLAHLKEOKNLSFPAVMFGLSNSA 726
 QY 181 ISRLAHTWERLPHKVRKLYSALERLDPSPNNHRYRLALAKLSPPIVFPMLLLKDMTFI 240
 DB 727 ISRLAHTWERLPHKVRKLYSALERLDPSPNNHRYRLALAKLSPPIVFPMLLLKDMTFI 786
 QY 241 HEGNHTLVENLINFEEKRMMAARAAMLHCRSHNPVPLSPRSRVSHLHEDSOVARISTC 300
 DB 787 HEGNHTLVENLINFEEKRMMAARAAMLHCRSHNPVPLSPRSRVSHLHEDSOVARISTC 846
 QY 301 SEOSTSRPASTWAVYQOLKVIDNORELSRLSRELEP 338
 DB 847 SEOSTSRPASTWAVYQOLKVIDNORELSRLSRELEP 884

RESULT 4
 AAY70968 ID AAY70968 standard; Protein: 849 AA.
 AC AAY70968;
 DT 09-AUG-2000 (first entry)
 XX
 DE Human Ras signalling pathway associated protein CAMP-GEFI.
 XX
 KM Ras signalling pathway; CAMP-GEFI; cyclic adenosine monophosphate;
 XX

KW		guanine nucleotide exchange factor; Rap1a; diagnosis; treatment;
XV		CAMP-GEF-associated disorder; drug; transgenic animal model;
KM	Ras-associated cancer; protein therapy; human.	
XX	Homo sapiens.	
OS		
XX		
FH	Key	Location/Qualifiers
FT	Binding-site	222..291
ET	/label= CAMP-binding_domain	
ET	/note= "amino acid at position 271 determines binding specificity for cAMP/cGMP"	
ET	606..629	
ET	Region	
ET	/label= SCR1	
ET	/note= "Structurally conserved region which is highly homologous to Ras-superfamily GEFs"	
ET	678..720	
ET	Region	
ET	/label= SCR2	
ET	/note= "Structurally conserved region which is highly homologous to Ras-superfamily GEFs"	
ET	757..778	
ET	/label= SCR3	
ET	/note= "Structurally conserved region which is highly homologous to Ras-superfamily GEFs"	
XX		
PN	WO200024768-A2.	
PD	04-MAY-2000.	
XX		
PF	22-OCT-1999;	99WO-US24826.
PR	23-OCT-1998;	98US-0105507.
PR	16-NOV-1998;	98US-0108685.
XX		
PA	(MASI) MASSACHUSETTS INST TECHNOLOGY.	
XX		
EI	Kawasaki H, Graybiel A, Hausman D;	
DR	WPI; 2000-350690/30.	
DR	N-PSDB; AAD00318.	
XX		
PT	Isolated nucleic acid comprises nucleotide sequence encoding protein selected from normal or mutant CalDAG-guanine nucleotide exchange factor -	
PS	Claim 64; Page 126-128; 128pp; English.	
CC	The present sequence is a human cyclic adenosine 3' , 5'	
CC	monophosphate-guanine nucleotide exchange factorII (cAMP-GEFII)	
CC	which has CAMP binding domain and Ras superfamily GEF	
CC	domains. It has substrate specificity for Rap1a and is differentially distributed in brain and various other tissues. CAMP-GEF proteins selectively activate the Ras superfamily small G protein and Rap1a, function as a negative regulator of Ras and directly couple the CAMP signal transduction system to Ras superfamily cascades. The present sequence is used for diagnosis and treatment (protein therapy) of CAMP-GEF-associated disorders. It is also useful for developing drugs and producing cell-lines or transgenic animal models for Ras-associated cancers.	
SO	Sequence	849 AA:
OY	Query Match	49.1%; Score 863.5; DB 21; Length 849;
Dn	Best Local Similarity	50.0%; Pred. No. 3.8e-78;
Matches	169; Conservative	63; Mismatches 81; Indels 25; Gaps 3
OY	1	MALAAQEGMTKRGVLVVNSAKDAIGLDPARGVATISGLNERLFVNPPQEVHELIPHP
Dn	534	ISAVALDKSGEGLIVYMSSGGEEVYLKPNDVSFFTYITINGRFACPREDFDSITLPD
OY	61	DOLGPVTASAGCLDVSAKDLAGOLTDDHWLSFNSHOVELHYVLGPGHLDDYTANLE
Dn	594	EDEPPTVGTVGTFFELMSKKDLAVQNTTIDWEFLNCVHEDELITYHFEG-RHNFRKTIANLD

QY	121	RFMRRELDGYWATELCGPPVGPAAOLRRKFETLLAHLKÖKKULNSFPVFMGLNSA	180
Db	653	LFERRENEIDFWWVTEICLCSQSLSKVÖLKKFFETIAAHCKEYKNLNSFFAIWGLS	712
QY	181	ISRLAHWEHLPRKVRKLYSALERLDDPSMNRVRYRLALAKLSPVPIPFMPLKDMTFI	240
Db	713	VSRLALWMEKLPKSKFKFFAAEFESLMDPSRNNHRAVRFLVAKLEPPLIPFPLKDMTF	772
QY	241	HEGNHTLVEVLIINFCKRMARARAPLHHCGRSHNPVPLSRVSHLHEDSQVARI	300
Db	773	HEGKKEFDIDLNVFEEKMRMTAMTARVRYRSQ---PPNPDAQAANKNHQDVR-----	822
QY	301	SEGLSTRSPASTWAVVOOLKVYDNORELSRI	338
Db	823	-----STVROLNVIDNORTLÖMSHRLEP	846
RESULT 5			
AAV70967			
ID	AAV70967	standard; Protein; 291 AA.	
XX	AAV70967;		
XX	09-AUG-2000	(first entry)	
DE	Rat Ras signalling pathway associated protein CAMP-GEFII.		
XX			
KM	Rat; Ras signalling pathway; CAMP-GEFII; cyclic adenosine monophosphate;		
KM	GEF; guanine nucleotide exchange factor; Rap1A; diagnosis; treatment;		
KW	CAMP-GEF-associated disorder: drug; transgenic animal model;		
XX	Ras-associated cancer; protein therapy.		
XX			
OS	Rattus norvegicus.		
XX			
FH	Key	Location/Qualifiers	
FT	Region	48..71	
FT		/label= SCR1	
FT		/note= "Structurally conserved region which is	
FT	Region	highly homologous to Ras-superfamily GEFs"	
FT		120..162	
FT		/label= SCR2	
FT		/note= "Structurally conserved region which is	
FT	Region	highly homologous to Ras-superfamily GEFs"	
FT		199..220	
FT		/label= SCR3	
FT		/note= "Structurally conserved region which is	
FT		highly homologous to Ras-superfamily GEFs"	
XX			
PN	W0200024768-A2.		
XX			
PD	04-MAY-2000.		
XX			
PE	22-OCT-1999;	99WO-US24826.	
XX			
XX	23-OCT-1998;	98US-0105507.	
PR	16-NOV-1998;	98US-0108685.	
XX			
PA	(MASI) MASSACHUSETTS INST TECHNOLOGY.		
XX			
PI	Kawasaki H, Graybiel A, Housman D;		
XX			
DR	WPI; 2000-350690/30.		
DR	N-PSDB; AAD00317.		
XX			
PT	Isolated nucleic acid comprises nucleotide sequence encoding protein		
PT	selected from normal or mutant CalDAG-guanine nucleotide exchange		
XX	factor -		
XX			
PS	Claim 64; Page 122; 128pp; English.		
XX			
CC	The present sequence is a rat cyclic adenosine 3', 5'		
CC	monophosphate-guanine nucleotide exchange factorII (CAMP-GEFII)		
CC	which has CAMP binding domain and Ras superfamily GEF		

CC domains. It has substrate specificity for Rap1a and is differentially
 CC distributed in brain and various other tissues. GMP-GEF proteins
 CC selectively activate the Ras superfamily small G protein and Rap1a,
 CC function as a negative regulator of Ras and directly couple the cAMP
 CC signal transduction system to Ras superfamily cascades. The present
 CC sequence is used for diagnosis and treatment (protein therapy) of
 CC cAMP-GEF-associated disorders. It is also useful for developing drugs
 CC and producing cell-lines or transgenic animal models for Ras-associated
 CC cancers.

XX
 SO Sequence 291 AA;

Query Match

44.7%; Score 786.5; DB 21; Length 291;

Best Local Similarity 50.2%; Pred. No. 5.3e-71;
 Matches 156; Conservative 51; Mismatches 79; Indels 25; Gaps 3;

QY 28 LGPDARVATSLGLENLEFVNPOEVHELIPHPDQLGPTVGSAGLDLVSAKDLAQLTD 87
 DB 3 LKNDVSVFTLTITNGRLFACPREQFSLPLPEOEPTGTGVTGTFELMSKDLAYOMT 62
 QY 88 HDKSLFNSIHQVELIHVLAGPQHLRDVTANTLERFMRPNELQYVWATELCIPVGPRA 147
 DB 63 YDMELEFVCVLELEIYITFG-RHNFKTTANLDLFLRPNEDIQVWVTEICLSQLSKRV 121
 QY 148 QLLRKFTKLAHLKEOKNLNSFFAVMGSLNSAISRLAHTWERLPHKVRKLYSALERLD 207
 DB 122 QLLKCKIKIAHCKEYKLNLSFGIYVGLSNVAESRLALTWMEKLPKSKFFVAFESLMD 181
 QY 208 PSNNHRYRLALAKLSPVIRPMPDLKDMTFIEGHNHTLVENLINEKRMARARML 267
 DB 182 PSNNHRYRLALAKLSPVIRPMPDLKDMTFIEGHNHTLVENLINEKRMARARML 241
 QY 268 HHCGRSHVPLSLRSHLHEDSOVARISTCSQSLSTRPASTWAVYQOQLVINDR 327
 DB 242 RYRISO---FPNDPAQAKNHQDVR-----SYRQLAVINDR 277
 QY 328 ELSRLSRELP 338
 DB 278 TLSOMSHRELP 288

RESULT 6

ABB59888 standard; Protein; 957 AA.

XX
 AC ABB59888;

XX
 DT 26-MAR-2002 (first entry)

XX
 DE Drosophila melanogaster polypeptide SEQ ID NO 6456.

XX
 KW Drosophila: developmental biology; cell signalling; insecticide;
 pharmaceutical.

XX
 OS Drosophila melanogaster.

XX
 PN W0200171042-A2.

XX
 PD 27-SEP-2001.

XX
 PF 23-MAR-2001; 2001WO-US09221.

XX
 PR 23-MAR-2000; 2000US-191637P.

XX
 PR 11-JUL-2000; 2000US-0614150.

XX
 PA (PEKE) PE CORP NY.

XX
 PI Venter JC, Adams M, LI PWD, Myers EW;

XX
 DR WPI; 2001-6556860/75.

XX
 DR N-PSDB; ABL03991.

XX
 PT New isolated nucleic acid detection reagent for detecting 1000 or more

PT genes from Drosophila and for elucidating cell signalling and cell-cell
 PT interactions -

XX
 PS Disclosure; SEQ ID NO 6456; 21pp + Sequence Listing; English.

XX
 CC The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (AB116176-AB130511), expressed DNA
 CC sequences (AB101840-AB116175) and the encoded proteins

CC
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.

XX
 SO Sequence 957 AA;

Query Match

39.6%; Score 696; DB 22; Length 957;

Best Local Similarity 44.4%; Pred. No. 4.3e-61;
 Matches 144; Conservative 54; Mismatches 100; Indels 26; Gaps 3;

QY 15 VLKVNAGDAIGLPDARGVATSLGNERLNVNPOEVHELIPHPDQLGPTVGSAGLD 74
 DB 656 VLVEKSNNGERSYFKDNDVSIPGTSLNGRLVSVKDLDALTLQOEQCEPTGVDIDLE 715
 QY 75 LYSAKDLAQLTDHDMSLFNSIHQVELIHVLAGPQHLRDVTANTLERFMRPNELQYVWA 134
 DB 716 ILSTKELAYHTLTTEWOLFVAVHELYELHYEGHNHGKI-TANLDVFLRFPNEVOYWI 774
 QY 135 TELCLCPVGPRAQLLKRFTKLAHLKEOKNLNSFFAVMGSLNSAISRLAHTWERLPHK 194
 DB 775 TELVSTPSLSKRVGLVYKFTKLAHYCKEYQNLNFFAVMGSLNSAISRLAHTWERLPHK 834
 QY 195 VRKLYSALERLDPSNMHRYRLALAKLSPVIRPMPDLKDMTFIEGHNHTLVENLINE 254
 DB 835 FRKIFQEFELIDPSRHNRAVRFVGLQPLIPFMPDLKDMTFIEGHNHTLVENLINE 894
 QY 255 EKMMARARARMLHHCGRSHVPLSLRSHLHEDSOVARISTCSQSLSTRPASTW 314
 DB 895 EKMMARARARMLHHCGRSHVPLSLRSHLHEDSOVARISTCSQSLSTRPASTW 929
 QY 315 AVYQOQLVINDRQLSRLSRELP 338
 DB 930 SYISSRVINDRQLVLTAMSQKEP 953

RESULT 7

AAW78706 standard; Protein; 456 AA.

XX
 AC AAW78706;

XX
 DT 06-NOV-2001 (first entry)

XX
 DE Human protein SEQ ID NO 1368.

XX
 KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;
 KW vaccine; peptide therapy; stem cell growth factor; hematopoiesis;
 KW tissue growth factor; immunomodulatory; cancer; leukemia;
 KW nervous system disorder; arthritis; inflammation.

XX
 OS Homo sapiens.

XX
 PN W0200157190-A2.

XX
 PD 09-AUG-2001.

XX
 PF 05-FEB-2001; 2001WO-US04098.

XX
 PR 03-FEB-2000; 2000US-0496914.

XX
 PR 27-APR-2000; 2000US-0560875.

PR 20-JUN-2000; 2000US-0598075.
PR 19-JUL-2000; 2000US-0620325.
PR 01-SEP-2000; 2000US-0654936.
PR 15-SEP-2000; 2000US-0663561.
PR 20-OCT-2000; 2000US-0693325.
PR 30-NOV-2000; 2000US-0728422.
XX (HYSE-) HYSEQ INC.
XX
XX Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y, Ma Y;
PI Zhao QA, Wang D, Wehrman T, Ren F, Chen R, Wang ZW;
PI Xue AJ, Yang Y, Wehrman T, Goodrich R,
XX
DR WPI; 2001-476283/51.
DR N-PSDB; AAK51839.
XX
PT Nucleic acids encoding polypeptides with cytokine-like activities,
PT useful in diagnosis and gene therapy -
XX
PS Claim 20; Page 3619-3620; 6221pp; English.
XX
XX The invention relates to polynucleotides (AAK51456-AAK53435) and the
CC encoded polypeptides (AAW8323-AAW80302) that exhibit activity elating to
CC cytokine, cell proliferation or cell differentiation or which may induce
CC production of other cytokines in other cell populations. The
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
CC peptide therapy. The polypeptides have various cytokine-like activities,
CC e.g. stem cell growth factor activity, haematopoiesis regulating
CC activity, tissue growth factor activity, immunomodulatory activity and
CC activity/inhibit activity and may be useful in the diagnosis and/or
CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
CC inflammation.
CC Note: Records for SEQ ID NO 2110 (AAK52581), 2111 (AAK52582) and 3666
CC (AAW80020) are omitted as the relevant pages from the sequence listing
CC were missing at the time of publication.
XX
XX Sequence 456 AA:
SQ
Query Match 36.2%; Score 637.5; DB 22; Length 456;
Best Local Similarity 42.0%; Pred. No. 1.3e-55;
Matches 136; Conservative 60; Mismatches 99; Indels 29; Gaps 6;
QY 15 VLKVNAGDAGLQPDAGVATSGLNRLFPVNPQEVNHELPHRDQGPVGSABGLD 74
DB 157 ILVAVSSGSEKVLQPTEDCVFTALGINSHLFACSTRDSYEAUVLPPEETQVSPGDE-TH 215
QY 75 LVSADKDLAQGLTDHMSLFNSIHQVELIHVYL-GPQHLRDVYTTANLERFRFRNELQYVW 133
DB 216 RVEPDEVANHHTAFHMFELRCVHELEFPDYVHGERGRRE--TANLELLQRCSEVTHW 273
QY 134 ATELCLCPVPPRAQDLRRFKILAAHLKQKLNSEFAVFGLSNSAISRLAHWTWERLPH 193
DB 274 ATEVLLCEAPGRRAQDLKFKIKIAALCKQNDLLSFYAVMGIDNAVSRLLRTWELKPG 333
QY 194 KYRKLYLSALERLDPSNMHRVRLAKLSPIVYPPMPLLDMTFTHGNGNTLVENLNL 253
DB 334 KRNKLFKEFENTLDPCKRNHKSRYEVLSKKKPPVPIPVPLLDLFTLHGSKTLVDGLVN 393
QY 254 FEKRMAMARAARMLHRCSHNPVLSPLRSRVSHLEDQVARIPTCSQSLSTRSPAST 313
DB 394 IERKHSVAAKRVTKIRYRS-RPLCLD-MEASERNHL-----QT 428
QY 314 WATVQQLKVIDNORELSRLSRELE 337
DB 429 KAYVROFOYIDNQNLLFELSYKLE 452

RESULT 8
AAW83061
ID AAW83061 standard; Protein; 456 AA.
AC
XX AAW83061;
XX

DT 22-OCT-2001 (first entry)
XX
XX Human polypeptide SEQ ID NO 2206.
DE
XX Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW peripheral nervous system; neuropathy; central nervous system; CNS;
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KW leukaemia.
XX
XX Homo sapiens.
OS
XX
XX WO200153312-A1.
PN
XX
PD 26-JUL-2001.
XX
XX 26-DEC-2000; 2000WO-US34263.
PF
XX
XX 21-JAN-2000; 2000US-0488725.
PR 25-APR-2000; 2000US-0552317.
PR 09-JUL-2000; 2000US-0598042.
PR 19-JUL-2000; 2000US-0620312.
PR 03-ANG-2000; 2000US-0653450.
PR 14-SEP-2000; 2000US-0662191.
PR 19-OCT-2000; 2000US-0693036.
PR 29-NOV-2000; 2000US-0727344.
XX
XX (HYSE-) HYSEQ INC.
PA
XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;
XX
XX WPI; 2001-442253/47.
DR N-PSDB; AAI58217.
DR
XX
PT Novel nucleic acids and polypeptides, useful for treating disorders
PT such as central nervous system injuries -
XX
XX Example 4; SEQ ID NO 2206; 10078pp; English.
XX
CC The invention relates to human nucleic acids (AAI57798-AAI61369) and
CC the encoded polypeptides (AAW8642-AAW42213) with nootropic,
CC immunosuppressant and cytostatic activity. The polynucleotides are useful
CC in gene therapy. A composition containing a polypeptide or polynucleotide
CC of the invention may be used to treat diseases of the peripheral nervous
CC system, such as peripheral nervous injuries, peripheral neuropathy and
CC localised neuropathies and central nervous system diseases, such as
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC utilisation of the activities such as: Immune system suppression,
CC activation/inhibit activity, chemotactic/chemokinetic activity, haemostatic
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC assays for receptor activity, arthritis and inflammation, leukaemia and
CC C.N.S disorders.
CC Note: The sequence data for this patent did not form part of the printed
CC specification.
XX
XX Sequence 456 AA:
SQ
Query Match 36.2%; Score 637.5; DB 22; Length 456;
Best Local Similarity 42.0%; Pred. No. 1.3e-55;
Matches 136; Conservative 60; Mismatches 99; Indels 29; Gaps 6;
QY 15 VLKVNAGDAGLQPDAGVATSGLNRLFPVNPQEVNHELPHRDQGPVGSABGLD 74
DB 157 ILVAVSSGSEKVLQPTEDCVFTALGINSHLFACSTRDSYEAUVLPPEETQVSPGDE-TH 215
QY 75 LVSADKDLAQGLTDHMSLFNSIHQVELIHVYL-GPQHLRDVYTTANLERFRFRNELQYVW 133
DB 216 RVEPDEVANHHTAFHMFELRCVHELEFPDYVHGERGRRE--TANLELLQRCSEVTHW 273

OY	134	ATELCICVPGGRADLLKREKILAHNLKOKKUNLSEFAWFGISNSAISRATWRLRH	193
Db	274	ATEVLICAPGRRADLLKREKILAHNLKOKKUNLSEFAWFGISNSAISRATWRLRH	333
OY	194	KVRKLSALERLRLDSWNRHVRRLAKLSPVIRPMPRLKDMFIRHSGNHLVENLIN	253
Db	334	KPKNLFRFENLTDPCRHNHKSRYEISKKMPVIRPMPRLKDMFIRHSGNHLVENLIN	393
OY	254	FEKRMARARARMLHHCNSHNPVPLSRSRVSHLHEDSQVARISTCSQSLSTRSPAST	313
Db	394	IEKLHSAVEKVTIRKKYRS-RECLLD-MEASPNHL-----QT	428
OY	314	WAYVOQLVINDQRELSTRSRELE	337
Db	429	KAYRQFOYIDNQLLELSTYKLE	452
RESULT 9			
AAE22105			
AC	AAE22105	standard: Protein; 246 AA.	
XX			
DT	25-JUL-2002	(first entry)	
XX			
DE	Human	48921 guanine-nucleotide exchange cell phorbol-ester consensus.	
XX			
KW	Human: GTP releasing factor; 48921; gene therapy; cytosstatic; virucide;		
KW	neuroprotective; vulnerrery; hepatotropic; cardiant; hypotensive; cancer;		
KW	inflammation; diabetes; cellular proliferative disorder; differentiative;		
KW	brain; liver; blood vessel; haematopoietic neoplastic; cerebrovascular;		
KW	leukemia; meningitis; prion disease; Parkinson's disease; Pick disease;		
KW	Alzheimer's disease; Huntington's disease; infection; hepatic injury;		
KW	Wilson's disease; haemochromatosis; Gaucher's disease; atherosclerosis;		
KW	tumour; hypertension; Kaposi sarcoma; chromosome mapping; tissue typing;		
KW	transgenic animal.		
XX			
OS	Homo sapiens.		
XX			
PN	WO200218577-A2.		
XX			
PD	07-MAR-2002.		
XX			
PF	27-AUG-2001; 2001WO-US26696.		
XX			
PR	30-AUG-2000; 2000US-228760P.		
XX			
PA	(MILL-) MILLENNIUM PHARM INC.		
XX			
P1	Curtis RAJ;		
XX			
DR	WPI: 2002-351708/38.		
XX			
PT	New isolated human GTP releasing factor-48921 polypeptide for treating		
XX	cancer, inflammation, diabetes and pathogenic invasion of host cells		
XX			
PS	Disclosure: Fig 4; 110pp; English.		
XX			
CC	The invention relates to an isolated human GTP releasing factor, termed		
CC	48921 protein. Human 48921 DNA and protein are useful as diagnostic and		
CC	therapeutic agents for preventing a disease or condition associated with		
CC	an aberrant or unwanted 48921 activity in a subject. Including cancer,		
CC	inflammation, diabetes and pathogenic invasion of host cells. 48921		
CC	protein is also useful for treating various disorders, including cellular		
CC	proliferative and differentiative disorders (metastatic, malignant,		
CC	cancer, hematopoietic neoplastic disorders including leukaemia), brain		
CC	(cerebrovascular disorders, meningitis, prion disease, Pick disease,		
CC	Alzheimer's disease, Parkinson's disease, Huntington's disease, inborn		
CC	errors of metabolism), liver (viral, bacterial or parasitic infections,		
CC	hepatic injury, Wilson's disease, haemochromatosis, Gaucher's disease),		
CC	blood vessel (atherosclerosis, tumours, hypertension, Kaposi sarcoma and		
CC	pathology of therapeutic interventions in vascular diseases). 48921 DNA		
CC	and protein are useful in screening assays, detection assays (forensic		

Query Match	23.8%	Score 419.5	DB 23	Length 246
Best Local Similarity	43.3%	Pred. No. 6,4e-34		
Matches	91	Conservative	40	Mismatches 58; Indels 21; Gaps 5
Oy	73	LDL--VSAKKDLAAGLTDDHMSLEFNGHVELTHYV-----LGPOLHLDVTANLERFMR	124	
Db	8	LDLHLDPEELAEQTLTDLDFELFKTEPSECCGIYWSNRKKGKENTL-----SPNLEKFIQ	63	
Oy	125	RFNELQYVATELCLCPVGPRAOILRRFKIKLAHLKEQKNLNSFFAVFGLSNSAIRL	184	
Db	64	RPNNTSYVAWEIILSESEKPDQAKIIEKFIVKAQCHRELNNFNSLMAIVSGLNSSSIYRL	123	
Oy	185	AHTWERLEHKKRKLKLSALERLLDPSNNHRYRLALAKLS-----PPYIPMPPLLLKDM	237	
Db	124	KKTWEKVPKREKRLFEELSELMDPSSNNYKNRYELLKSGTFVNSHOPPCIPFLGYLKL	183	
Oy	238	TFIHGNGHTIVE--NLINPEKKRMARAR	265	
Db	184	TFIHGNDYLDNTNLINPEKKRMIAKIIIR	213	
RESULT 10				
AAB07792				
ID	AAB07792	standard: Protein; 1499 AA.		
XX	AC	AAB07792;		
XX	DT	07-NOV-2000 (first entry)		
XX	DE	A murine guanine nucleotide releasing factor 4.		
XX	KW	guanine nucleotide releasing factor 4; GRF-4; Ras activator;		
KW	Nedd4 ubiquitination; cell metabolism; cell proliferation; cancer;			
XX	OS	cell differentiation; cell transformation; neuronal disorder.		
XX	Mus sp.			
XX	Key	Location/Qualifiers		
XX	FT	135..253		
FT	Domain	/note="cNMP-binding domain"		
FT	Domain	266..322		
FT	Domain	/note="REM domain"		
FT	Domain	386..470		
FT	Domain	/note="PDZ domain"		
FT	Domain	594..692		
FT	Domain	/note="RA domain"		
FT	Domain	712..899		
FT	Region	/note="CDC25 domain"		
FT	Region	1403..1406		
FT	Region	/note="PY motif"		
FT	Region	1425..1428		
XX	Region	/note="PY motif"		
XX	PN	WO20043510-A2.		
XX	PD	27-JUL-2000.		
XX	PE	20-JAN-2000; 2000WO-CA00042.		
PR	20-JAN-1999;	99CA-2259830.		

PA (HSCR-) HSC RES & DEV LP.
XX
XX
PI Rotin D, Pham N;
XX
DR WPI: 2000-499228/44.
XX N-PSDB; AAA59383.
XX
PT Nucleic acids encoding guanine nucleotide releasing factor-4 useful for
PT the treatment of cancers and neuronal disorders -
XX
PS Claim 9; Fig 19A; 89pp; English.
XX
CC The present sequence represents a murine guanine nucleotide releasing
CC factor (GRF)-4 (Ras activator) polypeptide. GRF4 activates Ras both
CC in vitro and in vivo. It directly binds cyclic adenosine monophosphate
CC (cAMP) directly via its CNMP-BD (cAMP/guanine monophosphate (cGMP)
CC binding domain). GRF4 directly connects cAMP-generating (e.g. G protein
CC coupled receptors) or cGMP-generating pathways to Ras. GRF4 activates
CC Ras in response to elevation of intracellular cAMP and/or cGMP. GRF4
CC is a target for Nedd4 ubiquitination as it binds Nedd4. Activation of
CC the Ras signalling pathway controls numerous cellular functions, such as
CC cell metabolism, proliferation, differentiation and transformation.
CC Therefore modulation of Ras activity may provide a mechanism for
CC controlling diseases. GRF4 polynucleotides and polypeptides may be used
CC in the treatment of diseases associated with inappropriate GRF4
CC expression and activity such as cancers and neuronal disorders. The
CC GRF4 polypeptides may be used as antigens in the production of
CC antibodies against GRF4 and in assays to identify modulators (agonists
CC and antagonists) of GRF4 expression and activity. The anti-GRF4
CC antibodies and GRF4 antagonists may also be used to down regulate GRF4
CC expression and activity. Inhibition of Ras can reduce cellulose
CC proliferation and cancers.
XX
SQ Sequence 1499 AA:
Query Match 23.3%; Score 409; DB 21; Length 1499;
Best Local Similarity 30.9%; Pred. No. 9.9e-32;
Matches 107; Conservative 57; Mismatches 122; Indels 60; Gaps 5;
QY 35 VATSLGILNRLFYVNPQEVHELIPHPDQLGPTVSAEGDLVSAKDLAQLTJHDWSLFN 94
DB 675 LADRILOSGRYVYKNNMFEPTLCSDDDAQELARESQISLQSTVEVAQVLSKRNELPR 734
QY 95 SIHOVELIHVYLGPOHLRDVTT-ANIEREMRFRNELQYVAVTELCCLCPVGPRAQLIKRF 153
DB 735 NIEPTXYIDLF--KLRKSTSCANKRFEVINOCTFWVASILLETQOLKRMKIKIHR 791
QY 154 IKAALHKBOKNLSNFFAVWFGLSAISRLATHTWRLPKVKYKLSALERLIDPSMNR 213
DB 792 IKIALCRBCKNNSMFAIISGLNLAIPVAKLTWTWELPKRYEKLRFQDLODLDFPSRMA 851
QY 214 VYRLAL--AKLSPVPIPFMPLLLKDMTFIEGHNHTLVENLNEKRMARARMLHGR 271
DB 852 KYRNVLNSQNLQPIIPLPVVIKKDLFLHEGNDSDVDGLVNEKRLMAKELRHGRMA 911
QY 272 SHNPVPLSPKRSR-----VSRLHEDS 292
DB 912 SYVMADMALMERTRKKRKRSLGSLDSQSTNAVYLDVAGTGKKRVRSSFLNAKKLYEDA 971
QY 293 QVARISTCSQSLSRSPASTWAVYQQLVINDQRELRLSPRELPR 338
DB 972 QMAK-----KVKOYLSNLLEMDSESLQTLISQCEP 1002
RESULT 11
AAE13102
ID AAE13102 standard; protein; 261 AA.
XX
AC AAE13102;
XX
DT 28-JAN-2002 (first entry)
XX
DE Human phospholipase C 16836 protein RasGER domain consensus sequence.

XX
KW Human; phospholipase C; PLC; 16836 protein; cellular disorder; sarcoma;
KW carcinoma; metastatic disorder; leukaemia; bone disorder; osteoporosis;
KW haematopoietic neoplastic disorder; sarcoidosis; Alzheimer's disease;
KW Parkinson's disease; brain disorder; reproductive disorder; gonorrhoea;
KW teratoma; immune disorder; rheumatoid arthritis; cardiovascular disorder;
KW Grave's disease; atherosclerosis; cardiomyopathy; anorexia nervosa;
KW obesity; cachexia; lipid disorder; diabetes; hepatocellular cancer;
KW viral disease; gene therapy; cytostatic; osteopathic; immunosuppressive;
KW neuroprotective; dermatological; ophthalmological; nootropic; vasotropic;
KW hypotensive; anticonvulsant; antibacterial; tuberculostatic; virucide;
KW veneral; anorectic.
XX
OS Homo sapiens.
XX
PN WO200175121-A2.
XX
XX 11-OCT-2001.
PD
XX 30-MAR-2001; 2001WO-US10273.
PF
XX 31-MAR-2000; 2000US-193921P.
PR
XX (MILL-) MILLENNIUM PHARM INC.
PA Meyers RA, Hunter JJ;
PI WPI: 2001-662975/76.
DR
XX
XX
XX
PT New polypeptide for controlling cellular proliferative and/or
PT differential disorders such as lung, breast or colon cancers,
PT comprises a member of human phospholipase C family, referred to as
PT 16836 -
XX
XX
PS Disclosure; Fig 3A; 148pp; English.
XX
XX The present invention relates to novel human phospholipase C (PLC)
CC family members, designated as 16836 and polynucleotides encoding
CC them. 16836 sequences are useful for treating cellular proliferative
CC and/or differentiative disorders such as carcinoma, sarcoma, metastatic
CC disorders or haematopoietic neoplastic disorders (e.g. leukaemia)
CC which arise from myeloid, lymphoid or erythroid lineages or their
CC precursor cells. They are also useful for treating bone disorders
CC (e.g. osteoporosis, sarcoidosis), brain disorders (e.g. Alzheimer's
CC disease, Parkinson's disease), reproductive disorders (e.g. gonorrhoea,
CC teratoma), immune disorders (e.g. rheumatoid arthritis, Grave's disease),
CC cardiovascular disorders (atherosclerosis, cardiomyopathies) or diseases
CC of metabolic imbalance such as obesity, anorexia nervosa, cachexia,
CC lipid disorders and diabetes. 16836 molecules of the invention are
CC used to treat disorders associated with an accumulation in the liver
CC of fibrous tissue, such as that resulting from an imbalance between
CC production and degradation of the extracellular matrix accompanied
CC by the collapse and condensation of preexisting fibers. They are
CC useful as markers for precursors or predisposition of disorders
CC or disease states, as markers of drug activity or as markers of the
CC pharmacogenomic profile of a subject. Modulators of 16836 activity
CC could be used to control viral diseases. They are used in the treatment
CC and/or diagnosis of virus-associated carcinoma, especially hepatocellular
CC cancer. 16836 sequences are also used in gene therapy. The present
CC sequence is human phospholipase C (PLC) 16836 protein Ras guanine
CC nucleotide exchange factor (RasGER) domain consensus sequence.
XX
SQ Sequence 261 AA:
Query Match 23.1%; Score 405.5; DB 22; Length 261;
Best Local Similarity 37.6%; Pred. No. 1.8e-32;
Matches 106; Conservative 41; Mismatches 92; Indels 43; Gaps 8;
QY 75 LVSAKDLAQLTJHDWSLNSIHQVELIHVYLGPOHLRD--VTANIEREMRFRNELQY 131
DB 3 LIDPELALQTLTLDPBELFRKIEPSLGSVWGKSKKSPPLAQNLQNEAFIERENEVS 62
QY 132 WYATELCDCPV--PGPRAQLKRFIKLAHLKBOKNLSNFFAVWFGLSAISRLAHWE 189

Db 63 WATTELCQTTLPKRAEVLSEFIYAKKCHRELNNFNSIMAIYASALSSSPISRLKKTWE 122
 Oy 190 RLPKRYKLYSALERLDPG-WNHRVYRLAL-----AKSPVYIPMPMLLDMPFIHE 242
 Db 123 KLSKTKKLPFEELLELDPSERNFKNYREALKSCNKPNOQPCVPPLGYLKDLPFIDE 182
 Oy 243 GNHTLVEN-----LINFEMKMMARAARMLHCRSH-NPYLSPLRKSVSHLHDSQVARI 297
 Db 183 GNPDLFENGTKGLVNEFKRKRIKILREIQLOSACOPYLKKNRNDIQELLRAS----- 237
 Oy 298 STCSEOSLSTRSPASTWAVYQOLKVI-DNORELSRLSRELEP 338
 Db 238 -----RPLEVLPEEDELVELSLRIEP 259

RESULT 12
 AAU99911
 ID AAU99911 standard; Protein: 261 AA.
 XX
 AC AAU99911;
 XX
 DF 07-OCT-2002 (first entry)
 XX
 DE Human guanine nucleotide dissociation stimulator consensus sequence #2.
 XX
 KW 47476: guanine nucleotide dissociation factor infection;
 KW haemotopoietic disorder; blood clotting disorder; cancer;
 KW autoimmune disorder; leukaemia; immunological disorder;
 KW cardiovascular disorder; neurological disorder; cellular proliferation;
 KW red blood cell disorder; viral disease; neurological disorder.
 XX
 OS Synthetic.
 XX
 PN WO200240656-A2.
 XX
 PD 23-MAY-2002.
 XX
 PF 14-NOV-2001; 2001WO-US45291.
 XX
 PR 14-NOV-2000; 2000US-248331P.
 PR 14-NOV-2000; 2000US-248362P.
 PR 14-NOV-2000; 2000US-248365P.
 PR 30-NOV-2000; 2000US-250077P.
 PR 30-NOV-2000; 2000US-250176P.
 PR 30-NOV-2000; 2000US-250327P.
 XX
 PA (MILL-) MILLENNIUM PHARM INC.
 XX
 PI Meyers RE, Curtis RAJ, Gluckmann MA;
 XX
 DR WPI; 2002-508325/54.
 XX
 XX Isolated 47476, 67210, 49875, 46842, 33201, 83378, 84233, 64708, 85041
 PT or 84234 polypeptides, useful as reagents or targets for treating or
 PT diagnosing pain or metabolic, liver, kidney, or cardiovascular
 PT disorders
 XX
 PS Disclosure; Fig 2; 298pp; English.

CC biosynthesis or glycogen synthesis) immunological disorders,
 CC cardiovascular disorders, neurological disorders, or cellular
 CC proliferation and/or differentiation disorders, e.g., cancer, cell
 CC motility and adhesion disorders, differential disorders, red blood
 CC cell disorders, viral diseases, neurological disorders (e.g., brain
 CC disorders), pain or metabolic disorders, liver disorders, kidney
 CC imbalance, protein trafficking disorders and disorders of metal ion
 CC bone metabolism. The sequences of the invention are also useful for
 CC screening assays, predictive medicine (e.g., diagnostic assays,
 CC prognostic assays, monitoring clinical trials, and pharmacogenetics);
 CC and methods of treatment (e.g., therapeutic and prophylactic). The
 CC present sequence represents a predicted consensus sequence motif
 CC found in the protein of the invention.

SO Sequence 261 AA;
 Query Match 23.1%; Score 405.5; DB 23; Length 261;
 Best Local Similarity 37.6%; Pred. No. 1.ee-32;
 Matches 106; Conservative 41; Mismatches 92; Indels 43; Gaps 8;

Oy 75 LVSAKDLAQGLTDHDMLEFNSIHVELIHVLPQHIRD---VTTALERFMRRNPLOY 131
 Db 3 LDPEELAEQTLTLDPELFRRKIEPSELGSGWKRKSKSPSPAPONLEAFIERFNEVSN 62
 Oy 132 WATTELCICPV--PGPRAQLRKFIKLAHLKEDKNLSPFAVWFGLSNISRLAHTWE 189
 Db 63 WATTELCQTTLPKRAEVLSEFIYAKKCHRELNNFNSIMAIYASALSSSPISRLKKTWE 122
 Oy 190 RLPKRYKLYSALERLDPG-WNHRVYRLAL-----AKSPVYIPMPMLLDMPFIHE 242
 Db 123 KLSKTKKLPFEELLELDPSERNFKNYREALKSCNKPNOQPCVPPLGYLKDLPFIDE 182
 Oy 243 GNHTLVEN-----LINFEMKMMARAARMLHCRSH-NPYLSPLRKSVSHLHDSQVARI 297
 Db 183 GNPDLFENGTKGLVNEFKRKRIKILREIQLOSACOPYLKKNRNDIQELLRAS----- 237
 Oy 298 STCSEOSLSTRSPASTWAVYQOLKVI-DNORELSRLSRELEP 338
 Db 238 -----RPLEVLPEEDELVELSLRIEP 259

RESULT 13
 ABB64459
 ID ABB64459 standard; Protein: 1551 AA.
 XX
 AC ABB64459;
 XX
 DT 26-MAR-2002 (first entry)
 XX
 DE Drosophila melanogaster polypeptide SEQ ID NO 20169.
 XX
 KW Drosophila; developmental biology; cell signalling; insecticide;
 KW pharmaceutical.
 OS Drosophila melanogaster.
 XX
 PN WO200171042-A2.
 XX
 PD 27-SEP-2001.
 XX
 PF 23-MAR-2001; 2001WO-US09231.
 PR 23-MAR-2000; 2000US-191637P.
 PR 11-JUL-2000; 2000US-0614150.
 XX
 PA (PEKE) PE CORP NY.
 XX
 PI Venter JC, Adams M, Li PWD, Myers EW;
 XX
 DR WPI; 2001-656860/75.
 DR N-PSDB; ABL08562.

PT New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
XX interactions -
PS
XX
PS Disclosure: SEQ ID NO 20169; 21bp + Sequence Listing; English.
XX
CC The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (AB10176-AB130511), expressed DNA
CC sequences (AB101840-AB16175) and the encoded proteins
CC (AB85737-AB872072).
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 1551 AA:

Query Match 22.9%; Score 402.5; DB 22; Length 1551;
Best Local Similarity 33.8%; Pred. No. 4.7e-31;
Matches 110; Conservative 53; Mismatches 111; Indels 51; Gaps 10;

QY 61 DOLGPVGSAGLDLV-----SAKDLAQLTDDHWSLFNSIHVELIHVLCPOH 110
DB 842 DSTEPLVPDELALTELVRESNVFELHMLAYELAIQILIDP---FANRQIESTEYDELFE 898
QY 111 LSD-VTTANLERPMRFRNFELQYVATFELCLCPVPRQALRKFKLAHLKEOKNLNF 169
DB 899 LMSRGVPMLSFAELVNRKMFVSVSEICAEHNIVRKMTIVQFKIARHCKEKNFNSM 958
QY 170 FAVMEGLNSAISRLAHTWERLPHKVRKLYSALERLLDPMSMHRVYR-DATAKL--SPV 226
DB 959 FAIVSGLGGAVSRLRQWTEKLPSKYQRLFNDLQDLPMSRMMSKYRQLVSAELLAQHPI 1018
QY 227 IFPMPLLKDMTFIHGNTLVENLINFKMMARMAAAMLHCHSHNPVL----- 278
DB 1019 IFPPYIKKDLTFIHGNTDVRDGLINFELKMLKEVLLTHMCS-SPYDLSTLELKG 1077
QY 279 -SPPLR-----SHVSHLEDQVARI-----STCSEOSTLRSP-----AST 313
DB 1078 QSPSNALFSLNOMASQSNAAAGTYIAANAGQATTKRRKASTAARPKMFEAQMVRV 1137
QY 314 MAYVOQLKYIDNORELSTLSRELEP 338
DB 1138 KAYLNSIKILSDDELTKFLECEP 1162

RESULT 14
ABG66725
ID ABG66725 standard; Protein; 1651 AA.
XX
AC ABG66725;
XX
DT 30-AUG-2002 (first entry)
XX
XX Human novel polypeptide #60.
XX
XX Human: inflammatory condition; shock; sepsis; immune response;
XX cancer; wound healing; central nervous system disease; haematopoiesis;
XX peripheral nervous system disease; amyotrophic lateral sclerosis; tendon;
XX myeloid cell disorder; lymphoid cell disorder; platelet disorder; bone;
XX cartilage; ligament; nerve tissue; ulcer; osteoporosis; osteoarthritis;
XX bone degenerative disorder; periodontal disease; reperfusion injury;
XX lung fibrosis; liver fibrosis; autoimmune disorder; bacterial infection;
XX allergic condition; thrombolysis; thrombosis; coagulation disorder;
XX fungal infection.
XX
XX Homo sapiens.
XX
XX WO200244340-A2.
XX
XX

PD 06-JUN-2002.
XX
XX 30-NOV-2001; 2001WO-US47004.
XX
XX 30-NOV-2000; 2000US-0028952.
XX
XX (HYSE-) HYSEQ INC.
XX
XX Tang YF, Goodrich RW, Liu C, Zhou P, Asundi V, Wang J, Wang D,
XX Yamazaki V, Ujwal ML, Drmanac RT;
XX
XX WPI: 2002-508509/54.
XX
XX N-PSDB: ABR94949.
XX
XX Novel nucleic acids and polypeptides for diagnosis, treatment of
XX inflammatory, autoimmune, nervous system, myeloid or lymphoid cell
XX disorders, cancer and promoting wound healing -
XX
XX Claim 10; Page 629-632; 672pp; English.
XX
XX The invention relates to human novel polynucleotides and associated
XX polypeptides. The polynucleotides and polypeptides are useful for
XX treating inflammatory conditions such as arthritis, nephritis, Crohn's
XX disease, ischaemia-reperfusion injury, shock, sepsis, immune responses
XX and cancer and for promoting wound healing. The sequences are used to
XX induce the proliferation of neural cells and regeneration of nerve and
XX brain tissue, and are useful for the treatment of central and peripheral
XX nervous system diseases and neuropathies, such as Alzheimer's disease,
XX Parkinson's disease, Huntington's disease and amyotrophic lateral
XX sclerosis. The sequences are involved in chemoattract or chemokinetic
XX activity, regulation of haematopoiesis, treatment of myeloid or lymphoid
XX cell disorders and platelet disorders such as thrombocytopenia,
XX regeneration of bone, cartilage, tendon, ligament and/or nerve tissue
XX growth, tissue repair, healing of burns, incisions, ulcers, treatment of
XX osteoporosis, osteoarthritis, bone degenerative disorders and periodontal
XX disease. The sequences of the invention are also useful for gut
XX protection or regeneration and treatment of lung or liver fibrosis,
XX reperfusion injury in various tissues, immune deficiencies and disorders
XX including severe combined immunodeficiency (SCID), bacterial or fungal
XX infections, autoimmune disorders e.g. multiple sclerosis and myasthenia
XX gravis, allergic conditions such as asthma, thrombocytosis or thrombosis
XX and coagulation disorders. Sequences ABG66666-ABG66756 represent human
XX novel polypeptides of the invention.
XX
SQ Sequence 1651 AA:

Query Match 22.2%; Score 390.5; DB 23; Length 1651;
Best Local Similarity 30.0%; Pred. No. 8.5e-30;
Matches 101; Conservative 57; Mismatches 128; Indels 51; Gaps 5;

QY 35 VATSLGLNERLEFVNVRQEVHELIPHPDQGLPTVGSAGEGLDVSADKDLAQLTDDHWSLFN 94
DB 868 IADRIQLNGRYVYLKNNMETETLCSDEDAQELVKEQSMLQSLSTEVATQLSMRFDLFR 927
QY 95 SIHOVELLHYVLGPHLADVT-TANLERPMRFRNFELQYVATFELCLCPVPRQALRKFE 153
DB 928 NIEPTEYIDDLF--KLMSKTGNTHLKRFEEDIVQDETGVASSELITTEANOLKRIIKHF 984
QY 154 IKLAHLKEOKNLNFSAVMEGLNSAISRLAHTWERLPHKVRKLYSALERLLDPMSMHR 213
DB 965 IKIALHCECKNFNSMFAIISGLNLAAYARLGRYTEKPSYKEKHLQDLPDRSRMA 1044
QY 214 VYR--LALAKSPPIVPMPLLKDMTFIHGNTLVENLINFKMMARMAAAMLHCHCR 271
DB 1045 KYRNILSSQSQMQPRTPLFPVYKDMTFLEHGNSDKVGLVNEFKLRMSKREIRQVYMT 1104
QY 272 SHNPVPLPLPSR-----VSHLEDQVARI----- 301
DB 1105 SANMDPAMMFQROSLSSQSTNSNMLDVGGAAHKRRARRSSLLNAKKIYEDQAMAR----- 1159
QY 302 EQSLSTRSPASTWAVVQOLKVIDNORELSTLSRELEP 338
DB 1160 -----KVQYLSLSDVETDEKFMQMSLOWEP 1166

RESULT 15
 AAB42658
 ID AAB42658 standard: Protein: 1675 AA.
 XX
 AC AAB42658;
 XX
 DT 08-FEB-2001 (first entry)
 DE Human ORFX ORF2422 polypeptide sequence SEQ ID NO:4844.
 XX
 KW Human: Open reading frame: ORFX; detection; cytosolic; hepatotropic;
 KW vulnary; antiparasitic; antiparkinsonian; noctropic; neuroprotective;
 KW anticonvulsant; osteopathic; antilethritic; immunosuppressant; cardiant;
 KW immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;
 KW hypotensive; dermatological; immunosuppressive; antilethritic;
 KW antiviral; antibacterial; antifungal; antineumatic; antihypertensive;
 KW antianemic; gene therapy; cancer; proliferative disorder; hypertension;
 KW neurodegenerative disorder; osteoarthritis; graft vs host disease;
 KW cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;
 KW cholesterol ester storage; systemic lupus erythematosus; infection;
 KW severe combined immunodeficiency; malaria; autoimmune disorder; asthma;
 KW allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;
 KW bone damage; cartilage damage; antiinflammatory disease; coagulation;
 KW thrombosis; contraceptive.
 XX
 OS Homo sapiens.
 PN WO200058473-A2.
 PD 05-OCT-2000.
 XX
 PF 31-MAR-2000; 2000MO-US08621.
 XX
 PR 31-MAR-1999; 99US-0127607.
 PR 02-APR-1999; 99US-0127636.
 PR 05-APR-1999; 99US-0127728.
 PR 30-MAR-2000; 2000US-0540763.
 XX
 PA (CURA-) CURAGEN CORP.
 XX
 PI Shinkets RA, Leach M;
 XX
 DR WPI: 2000-602362/57.
 DR N-PSDB; AAC76867.
 XX
 PT Novel nucleic acids and peptides derived from open reading frame X,
 PT useful for treating e.g. cancers, proliferative disorders,
 PT neurodegenerative disorders and cardiovascular disease -
 XX
 PS Claim 11: Page 4028-4032; 5507pp; English.
 CC AAC74446 to AAC77606 encode the proteins given in AAB40237 to AAB43397,
 CC which represent the human ORFX open reading frames 1 to 3161. The ORFX
 CC sequences have activities such as: cytostatic; hepatotropic; vulnary;
 CC antiparasitic; antiparkinsonian; noctropic; neuroprotective;
 CC osteopathic; anticonvulsant; antilethritic; immunosuppressant;
 CC immunostimulant; cardiant; thrombolytic; coagulant; vasotropic;
 CC antidiabetic; hypotensive; dermatological; immunosuppressive;
 CC antilethritic; antibacterial; antifungal; antineumatic;
 CC antihypertensive; antianemic. The sequences can be used for determining
 CC the presence of or predisposition to, or preventing or treating
 CC pathological conditions associated with an ORFX-associated disorder. The
 CC nucleic acids can be used to express ORFX proteins in gene therapy
 CC vectors. The proteins and nucleic acids may be used to treat cancers,
 CC proliferative disorders, neurodegenerative disorders, osteoarthritis,
 CC graft vs host disease, cardiovascular disease, diabetes mellitus,
 CC hypertension, hypothyroidism, cholesterol ester storage, systemic lupus
 CC erythematosus, severe combined immunodeficiency (SCID), AIDS, viral,
 CC bacterial or fungal infection, malaria, autoimmune disorders, asthma,
 CC allergies, aplastic anaemia, burns, wounds, bone and cartilage damage,
 CC nocturnal haemoglobinuria, antiinflammatory disease; to enhance

CC coagulation; to inhibit thrombosis; and as a contraceptive.
 XX
 SQ Sequence 1675 AA;
 Query Match 22.28; Score 390.5; DB 21; Length 1675;
 Best Local Similarity 30.0%; Pred. No. 8,7e-30;
 Matches 101; Conservative 57; Mismatches 128; Indels 51; Gaps 5;
 QY 35 VATSIGNEELFVNNQVEVHELPHRQQLPTGSAEGLDLVSAKDLQGLTDHMSLFN 94
 DB 891 LADRIQNGRYVYKNNMETETLCSDEDAQLVESQLSMLSTIEVATQLSMRDFLFR 950
 QY 95 SIHQVELIHYVLGPQHLRDVT-TANLERFMRRFNELOQVYVATELCQPVPGPRAQLRKF 153
 DB 951 NIEPTVIDDLF---KLNKSTGTMHLKRFEDIYQETFWVASLITFANOLKMKIKHF 1007
 QY 154 IKLAHLKEOKNLNSFPVAVFGLSNSAISRLAHTWERLPHKVKRLYSALERLLDPSVNR 213
 DB 1008 IKTALHCRECKNFSMFAISGLNLASVARLRGTWEKLPSEKYEKHLQDIDFDPNRMA 1067
 QY 214 VYR--TALAKLSPVTPFMPDLLKDMFTIEGNTLVENLINFEMRMARAARMLHCR 271
 DB 1068 KYRNILSSQSMQPIPLFPVVKDMFTLHGDSKVDGLVNEKLMISKEIRQVVRMT 1127
 QY 272 SHNPVPLSPLRSR-----VSHLHEDSQVARISTCS 301
 DB 1128 SANMDPAMFRRQSLSGSTNSMMLDVGGAHKKRARSSLLNAKKRIEDAQMAR----- 1182
 QY 302 EQSLSTRSPASTWAVYQQLKVIDNORELSRLSRELP 338
 DB 1183 -----KVKQYLSLDVETDERKQMSLQWER 1209

Search completed: October 7, 2003, 07:13:10
 Job Time : 88 secs

THIS PAGE BLANK (USPTO)